

DR	PROSITE; PS50011; PROTEIN KINASE DOM; 1.
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW	Alternative splicing; ATP-binding; Nuclear protein; Phosphorylation;
DW	Serine/threonine-protein kinase; Transferase; Tyrosine-protein kinase.
FT	DOMAIN 159 475 Protein kinase.
FT	NP BIND 165 173 ATP (By similarity).
FT	BINDING 189 189 ATP (By similarity).
FT	ACT SITE 286 Proton acceptor (By similarity).
FT	VARSPLIC 1 180 Missing (in isoform 2). /FTID=VSP_008205.
FT	MUTAGEN 189 189 K->R; Loss of function.
FT	CONFLICT 58 58 E -> Q (in Ref. 4).
FT	CONFLICT 141 141 D -> N (in Ref. 4).
FT	CONFLICT 344 344 R -> K (in Ref. 2).
SQ	SEQUENCE 481 AA; 57344 MW; F2C5G965900C12AA CRC64;
Query Match	93.8%; Score 2261.5; DB 1; Length 481;
Best Local Similarity	93.5%; Pred No. 5.1e-132;
Matches 420; Conservative	8; Mismatches 10; Indels 11; Gaps 2;
QY	8 SH-SVEEDTH-----PSHYLEARSLNERYDRDYRDYRYDEYNDYCEGVPRHYHRD 56 :
DB	33 SHSSQTENRKCHPHQFQSDCHYLEARCLNERYDRDYRYDEYNDYCEGVPRHYHRD 92
QY	57 IESGYRIHCSSKSVSRSSSPKRKNRHCCSHOSKSKSHRRKRSIIEDEEGHLICQS 116 : :
DB	93 VESTYRIHCSSKSVSRSSSPKRKNRPCASHOSKSKSHRRKRSIIEDEEGHLICQS 152
QY	117 DVLYARYEIVDTLGGAGFGKVCECDHGMDGMHVAVKIVGNVGRYREAAARSIQVLH 176 :
DB	153 DVLYARYEIVDTLGGAGFGKVCECDHGMDGLHFAVKIVGNVGRYREAAARSIQVLH 212
QY	177 STDPNVSFRVCQMLEWFHDHGHVCIVFELLGLSTYDFIKENSFLPFQIDHIROMAYOIQ 236 :
DB	213 STDPNVSFRVCQMLEWFHDHGHVCIVFELLGLSTYDFIKENSFLPFQIDHIROMAYOIQ 272
QY	237 SINFLHNKLHTDLKPENILFKVSDYVVKNSKKRDERITKNTDIKVDPGSATYDDE 296 :
DB	273 SINFLHNKLHTDLKPENILFKVSDYVVKNSKKRDERITKNTDIKVDPGSATYDDE 332
QY	297 HSTILVSTRHYRAPVILALGWSQPCDVMSIGCILLEYLGLTVPOTHDSEKHLAMMERI 356 :
DB	333 HSTILVSTRHYRAPVILALGWSQPCDVMSIGCILLEYLGLTVPOTHDSEKHLAMMERI 392
QY	357 LGPIPQHMIQTKRKKYFHNNOLDWEHSSAGRYVRRCCKPLKEFMLCHDEHEKLFDLV 416 :
DB	393 LGPIPAHMIQTKRKKYFHNNOLDWEHSSAGRYVRRCCKPLKEFMLCHDEHEKLFDLV 452
QY	417 RMLEYDPQRTITLDEALQHPFDLLKKX 445 :
DB	453 RMLEYDTPARRITLDEALQHPFDLLKKK 481
RESULT 3	
QN5V8	PRELIMINARY; PRT; 484 AA.
ID	QN5V8
AC	Q8NSV8
DT	01-OCT-2002 (TreeBLrel. 22, Created)
DT	01-OCT-2002 (TreeBLrel. 22, Last sequence update)
DT	01-WAR-2004 (TreeBLrel. 26, Last annotation update)
DE	CDC-like kinase 1.
OS	Name=CLK1;
GN	OS Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX	NCBI_TaxId=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Bone;
EX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA	Klausner R.D., Collins P.S., Wagner L., Shenmen C.F., Schuler G.D.,
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.W., Bhat N.K.,

This Page Blank (uspto)

QY 439 FDLKKK 445
 Db 421 FDLKKK 427

RESULT 9

US-10-801-671-4
 ; Sequence 4, Application US/10801671
 ; Publication No. US20040152123A1
 ; GENERAL INFORMATION:
 ; APPLICANT: YAN, Chunhua et al.
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; FILE REFERENCE: THERBOF
 ; CURRENT APPLICATION NUMBER: US/10/801,671
 ; PRIOR FILING DATE: 2004-03-17
 ; PRIOR APPLICATION NUMBER: 60/227,470
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: 09/810,671
 ; PRIOR FILING DATE: 2001-03-19
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 427
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-801-671-4

Query Match 95.9%; Score 2312; DB 16; Length 427;
 Best Local Similarity 100.0%; Pred. No. 5.8e-170;
 Matches 427; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 HYLEARSLNERDYDRRYDEYNDYCEGVPRHYHDIESGVRIHCSKSSVRSRSPK 78
 Db 1 HYLEARSLNERDYDRRYDEYNDYCEGVPRHYHDIESGVRIHCSKSSVRSRSPK 60

QY 79 KRNRCSSHQSKSRSHRKRKRSIEDDEGHLCQSGDVLRARYEIVDTLGEAGFKV 138
 Db 61 KRNRCSSHQSKSRSHRKRKRSIEDDEGHLCQSGDVLRARYEIVDTLGEAGFKV 120

QY 139 ECIDHGMGMHVAVKVQNGVRYREARSEIOVLEHNLSTDNSVRCVQMLEWFDHGH 198
 Db 121 ECIDHGMGMHVAVKVQNGVRYREARSEIOVLEHNLSTDNSVRCVQMLEWFDHGH 180

QY 199 VCIVFELLGLSTYDFIKENSFLPFQIDHROMAYQICQSFNLFHNKLTHTDLKPNILF 258
 Db 181 VCIVFELLGLSTYDFIKENSFLPFQIDHROMAYQICQSFNLFHNKLTHTDLKPNILF 240

QY 259 VKSDYVVKYNSKGRDERTLKNITDKVDFGSAFYDDEHSTLVSTRHYRAPEVILALGW 318
 Db 241 VKSDYVVKYNSKGRDERTLKNITDKVDFGSAFYDDEHSTLVSTRHYRAPEVILALGW 300

QY 319 SOPCDVWSIGCILIEYLGFTVQTHDSKEHLAMMERILGPQHMIOKTRKRYFHHNQ 378
 Db 301 SOPCDVWSIGCILIEYLGFTVQTHDSKEHLAMMERILGPQHMIOKTRKRYFHHNQ 360

QY 379 LWDDEHSSAGRYVRRCKPLKFMCHDEHEKLFDLVRMLEYDPTQRTITLDEALQHPF 438
 Db 361 LWDDEHSSAGRYVRRCKPLKFMCHDEHEKLFDLVRMLEYDPTQRTITLDEALQHPF 420

QY 439 FDLKKK 445
 Db 421 FDLKKK 427

RESULT 10

US-09-905-999-25
 ; Sequence 25, Application US/09905999
 ; Patent No. US20020106771A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ULIRICH, Axel
 ; APPLICANT: NAYLER, Oliver

; TITLE OF INVENTION: CLK PROTEIN KINASES AND RELATED PRODUCTS AND METHODS
 ; FILE REFERENCE: 038602/0431
 ; CURRENT APPLICATION NUMBER: US/09/905,999
 ; CURRENT FILING DATE: 2001-07-17
 ; PRIOR APPLICATION NUMBER: 09/127,248
 ; PRIOR FILING DATE: 1999-07-31
 ; PRIOR APPLICATION NUMBER: PCT/IB97/00946
 ; PRIOR FILING DATE: 1997-06-17
 ; PRIOR APPLICATION NUMBER: US 08/877,150
 ; PRIOR FILING DATE: 1997-06-17
 ; PRIOR APPLICATION NUMBER: US 60/034,286
 ; PRIOR FILING DATE: 1996-12-19
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 25
 ; LENGTH: 481
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-09-905-999-25

Query Match 93.8%; Score 2261.5; DB 9; Length 481;
 Best Local Similarity 93.5%; Pred. No. 5.2e-166;
 Matches 420; Conservative 8; Mismatches 10; Indels 11; Gaps 2;

QY 8 SH-SVEEDTH-----PSHYLEARSLSNERDYDRRYDEYNDYCEGVPRHYHDI 56
 Db 33 SHSSTQENRCHCKPHQPKDSDCHYLEARCLNERDYDRRYDEYNDYCEGVPRHYHDI 92

QY 57 TESGYRIHCSKSSVRSRSPKRNCRSHQSKSRSHRKRKRSIEDDEGHLCQSG 116
 Db 93 VESTYRIHCSKSSVRSRSPKRNCRSHQSKSRSHRKRKRSIEDDEGHLCQSG 152

QY 117 DVLRARYEIVDTLGEAGFKVVECIDHGMGMHVAVKVQNGVRYREARSEIOVLEHNL 176
 Db 153 DVLRARYEIVDTLGEAGFKVVECIDHGMGMHVAVKVQNGVRYREARSEIOVLEHNL 212

QY 177 STDPSNVRVCVQMLEWFDHGHVCIIVFELLGLSTYDFIKENSFLPFQIDHROMAYQICQ 236
 Db 213 STDPSNVRVCVQMLEWFDHGHVCIIVFELLGLSTYDFIKENSFLPFQIDHROMAYQICQ 272

QY 237 SINFLHNKLTHTDLKPNILFVKSDYVVKYNSKGRDERTLKNITDKVDFGSAFYDDE 296
 Db 273 SINFLHNKLTHTDLKPNILFVKSDYVVKYNSKGRDERTLKNITDKVDFGSAFYDDE 332

QY 297 HHSTLVSTRHYRAPEVILALGWSQPCDVWSIGCILIEYLGFTVQTHDSKEHLAMMERI 356
 Db 333 HHSTLVSTRHYRAPEVILALGWSQPCDVWSIGCILIEYLGFTVQTHDSKEHLAMMERI 392

QY 357 LGPIPOHMIQKTRKRYFHHNQDDEHSSAGRYVRRCKPLKFMCHDEHEKLFDLV 416
 Db 393 LGPIPAHMIQKTRKRYFHHNQDDEHSSAGRYVRRCKPLKFMCHDEHEKLFDLV 452

QY 417 RRMLEYDPTQRTITLDEALQHPFDLKKK 445
 Db 453 RRMLEYDPTQRTITLDEALQHPFDLKKK 481

RESULT 11
 US-10-267-355
 ; Sequence 355, Application US/10267502
 ; Publication No. US20040071700A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kim, Jaeseob
 ; APPLICANT: Galant, Ron
 ; TITLE OF INVENTION: Obesity Linked Genes
 ; FILE REFERENCE: LSD-07416
 ; CURRENT APPLICATION NUMBER: US/10/267,502
 ; CURRENT FILING DATE: 2003-01-27
 ; NUMBER OF SEQ ID NOS: 439
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 355
 ; LENGTH: 481
 ; TYPE: PRT

This Page Blank (uspto)

This Page Blank (uspto)

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2005, 00:08:08 ; Search time 138 Seconds
(without alignments)
1063.630 Million cell updates/sec

Title: US-10-801-671-2

Perfect score: 2410

Sequence: 1 MCIPLEASHSVEDTHPSHY.....QRITLDEALQHPPFDLLKXK 445

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1396920 seqs, 329844858 residues

Total number of hits satisfying chosen parameters: 1396920

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

- Published Applications AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2410	100.0	445	9	US-09-810-671-2
2	2410	100.0	445	13	US-10-109-854-2
3	2410	100.0	445	14	US-10-339-656-2
4	2410	100.0	445	16	US-10-801-671-2
5	2315.5	96.1	481	15	US-10-267-502-352
6	2312	95.9	427	9	US-09-810-671-4
7	2312	95.9	427	13	US-10-109-854-4
8	2312	95.9	427	14	US-10-339-656-4
9	2312	95.9	427	16	US-10-801-671-4
10	2261.5	93.8	481	9	US-09-905-999-25
11	2261.5	93.8	481	15	US-10-267-502-355
12	1893	78.5	484	15	US-10-267-502-353
13	1887	78.3	429	9	US-09-810-671-5

14	1887	78.3	429	13	US-10-109-854-5	Sequence 5, Appli
15	1887	78.3	429	14	US-10-339-656-5	Sequence 5, Appli
16	1887	78.3	429	16	US-10-801-671-5	Sequence 5, Appli
17	1887	78.3	484	15	US-10-116-275-127	Sequence 127, App
18	1887	78.3	484	16	US-10-755-889-2	Sequence 2, Appli
19	1836	76.2	483	15	US-10-267-502-356	Sequence 356, App
20	1824	75.7	483	9	US-09-905-999-20	Sequence 20, Appli
21	1740	72.2	352	15	US-10-425-114-54366	Sequence 54366, A
22	1740	72.2	352	15	US-10-425-114-54510	Sequence 54510, A
23	1692	70.2	374	16	US-10-664-421-131	Sequence 131, App
24	1607	66.7	301	15	US-10-267-502-357	Sequence 357, App
25	1526	63.3	350	9	US-09-925-298-539	Sequence 539, App
26	1526	63.3	350	14	US-10-102-806-539	Sequence 539, App
27	1427.5	59.2	499	9	US-09-905-999-21	Sequence 21, Appli
28	1417	58.8	499	10	US-09-790-852-3	Sequence 3, Appli
29	1322	54.9	638	15	US-10-104-047-2626	Sequence 2626, Ap
30	1317	54.6	490	15	US-10-267-502-351	Sequence 351, App
31	1310.5	54.4	431	15	US-10-182-243-48	Sequence 48, Appli
32	1310	54.4	490	9	US-09-905-999-23	Sequence 23, Appli
33	1310	54.4	490	15	US-10-267-502-354	Sequence 354, App
34	1296	53.8	484	15	US-10-425-114-56866	Sequence 56866, A
35	1252	52.0	341	15	US-10-267-502-350	Sequence 350, App
36	1246	51.7	417	15	US-10-108-260A-4699	Sequence 4699, Ap
37	1222	50.7	511	15	US-10-267-502-349	Sequence 349, App
38	1222	50.7	517	13	US-10-108-605-135	Sequence 135, App
39	859	35.6	212	15	US-10-425-114-54148	Sequence 54148, A
40	848	35.2	242	15	US-10-425-114-53826	Sequence 53826, A
41	831.5	34.5	390	15	US-10-425-114-52391	Sequence 52391, A
42	821.5	34.1	429	15	US-10-425-114-55309	Sequence 55309, A
43	814.5	33.8	358	15	US-10-425-114-45337	Sequence 45337, A
44	802	33.3	543	16	US-10-437-963-179828	Sequence 179828, A
45	795	33.0	435	15	US-10-425-114-67650	Sequence 67650, A

ALIGNMENTS

RESULT 1

US-09-810-671-2
; Sequence 2, Application US/09810671
; Publication No. US20020076783A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758
; CURRENT APPLICATION NUMBER: US/09/810,671
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Human
US-09-810-671-2

Query Match	100.0%	Score	2410;	DB	9;	Length	445;
Best Local Similarity	100.0%	Pred. No.	1.7e-177;	Indels	0;	Gaps	0;
Matches	445;	Conservative	0;	Mismatches	0;		
Qy	1	MCIPLEASHSVEDTHPSHYLEARNLDRYDRRYDEYNDYCEGYVPRHYRDI	ESG	60			
Db	1	MCIPLEASHSVEDTHPSHYLEARNLDRYDRRYDEYNDYCEGYVPRHYRDI	ESG	60			
Qy	61	YRIHCSSSVRRSSSPKRNKNCSSHQSKSHRRKRSIEDDEGHLCQSGDVL	R	120			
Db	61	YRIHCSSSVRRSSSPKRNKNCSSHQSKSHRRKRSIEDDEGHLCQSGDVL	R	120			
Qy	121	ARYEIVDTLGGAGKVVCEIDHGMGHVAVKIVNNGRYREARSIQVLEHLS	TDP	180			
Db	121	ARYEIVDTLGGAGKVVCEIDHGMGHVAVKIVNNGRYREARSIQVLEHLS	TDP	180			

```
QY 181 NSVFCVQMLEWFDHGHVCIIVFELLGLSTYDFIKENSFLPQIDHIRMAYQICQSNF 240
DB 181 NSVFCVQMLEWFDHGHVCIIVFELLGLSTYDFIKENSFLPQIDHIRMAYQICQSNF 240
QY 241 LHNKLTHTDLKPENILFVKSIVVYKNSKMRDERTLKNITDIKVVDGSGATYDDEHST 300
DB 241 LHNKLTHTDLKPENILFVKSIVVYKNSKMRDERTLKNITDIKVVDGSGATYDDEHST 300
QY 301 LVSTRHYRAPEVILALGWSQPCDWSIGCILIEYLGTVFTQTHDSKEHLAMMERILGPI 360
DB 301 LVSTRHYRAPEVILALGWSQPCDWSIGCILIEYLGTVFTQTHDSKEHLAMMERILGPI 360
QY 361 PQHMIQKTRKRYFHHNQLDWEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLVRRML 420
DB 361 PQHMIQKTRKRYFHHNQLDWEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLVRRML 420
QY 421 EYDPTQRTITLDEALQHPFFDLKKK 445
DB 421 EYDPTQRTITLDEALQHPFFDLKKK 445

RESULT 2
US-10-109-854-2
; Sequence 2, Application US/10109854
; Publication No. US20020119548A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO00758DIV
; CURRENT APPLICATION NUMBER: US/10/109,854
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-109-854-2

Query Match 100.0%; Score 2410; DB 13; Length 445;
Best Local Similarity 100.0%; Pred. No. 1.7e-177;
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCIPLEASHSVEEDTHPSHYLEARSLSNERDYDRRYVDEYRNDYCEGVVPRHYRDIESG 60
DB 1 MCIPLEASHSVEEDTHPSHYLEARSLSNERDYDRRYVDEYRNDYCEGVVPRHYRDIESG 60
QY 61 YRIHCKSSVRSRSPKRNHCHSHQSRKSHRRKRSIEDDEGHILICQSGDVL 120
DB 61 YRIHCKSSVRSRSPKRNHCHSHQSRKSHRRKRSIEDDEGHILICQSGDVL 120
QY 121 ARYEIVDTLGEAGKVEICIDHGMGMHVAVKIVKNGRYREARSEIQVLEHLNSTDP 180
DB 121 ARYEIVDTLGEAGKVEICIDHGMGMHVAVKIVKNGRYREARSEIQVLEHLNSTDP 180
QY 181 NSVFCVQMLEWFDHGHVCIIVFELLGLSTYDFIKENSFLPQIDHIRMAYQICQSNF 240
DB 181 NSVFCVQMLEWFDHGHVCIIVFELLGLSTYDFIKENSFLPQIDHIRMAYQICQSNF 240
QY 241 LHNKLTHTDLKPENILFVKSIVVYKNSKMRDERTLKNITDIKVVDGSGATYDDEHST 300
DB 241 LHNKLTHTDLKPENILFVKSIVVYKNSKMRDERTLKNITDIKVVDGSGATYDDEHST 300
QY 301 LVSTRHYRAPEVILALGWSQPCDWSIGCILIEYLGTVFTQTHDSKEHLAMMERILGPI 360
DB 301 LVSTRHYRAPEVILALGWSQPCDWSIGCILIEYLGTVFTQTHDSKEHLAMMERILGPI 360
QY 361 PQHMIQKTRKRYFHHNQLDWEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLVRRML 420
DB 361 PQHMIQKTRKRYFHHNQLDWEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLVRRML 420
QY 421 EYDPTQRTITLDEALQHPFFDLKKK 445
DB 421 EYDPTQRTITLDEALQHPFFDLKKK 445

RESULT 3
US-10-339-656-2
; Sequence 2, Application US/10339656
; Publication No. US20030134319A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO00758DIV2
; CURRENT APPLICATION NUMBER: US/10/339,656
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/109,854
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-339-656-2

Query Match 100.0%; Score 2410; DB 14; Length 445;
Best Local Similarity 100.0%; Pred. No. 1.7e-177;
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCIPLEASHSVEEDTHPSHYLEARSLSNERDYDRRYVDEYRNDYCEGVVPRHYRDIESG 60
DB 1 MCIPLEASHSVEEDTHPSHYLEARSLSNERDYDRRYVDEYRNDYCEGVVPRHYRDIESG 60
QY 61 YRIHCKSSVRSRSPKRNHCHSHQSRKSHRRKRSIEDDEGHILICQSGDVL 120
DB 61 YRIHCKSSVRSRSPKRNHCHSHQSRKSHRRKRSIEDDEGHILICQSGDVL 120
QY 121 ARYEIVDTLGEAGKVEICIDHGMGMHVAVKIVKNGRYREARSEIQVLEHLNSTDP 180
DB 121 ARYEIVDTLGEAGKVEICIDHGMGMHVAVKIVKNGRYREARSEIQVLEHLNSTDP 180
QY 181 NSVFCVQMLEWFDHGHVCIIVFELLGLSTYDFIKENSFLPQIDHIRMAYQICQSNF 240
DB 181 NSVFCVQMLEWFDHGHVCIIVFELLGLSTYDFIKENSFLPQIDHIRMAYQICQSNF 240
QY 241 LHNKLTHTDLKPENILFVKSIVVYKNSKMRDERTLKNITDIKVVDGSGATYDDEHST 300
DB 241 LHNKLTHTDLKPENILFVKSIVVYKNSKMRDERTLKNITDIKVVDGSGATYDDEHST 300
QY 301 LVSTRHYRAPEVILALGWSQPCDWSIGCILIEYLGTVFTQTHDSKEHLAMMERILGPI 360
DB 301 LVSTRHYRAPEVILALGWSQPCDWSIGCILIEYLGTVFTQTHDSKEHLAMMERILGPI 360
QY 361 PQHMIQKTRKRYFHHNQLDWEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLVRRML 420
DB 361 PQHMIQKTRKRYFHHNQLDWEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLVRRML 420
QY 421 EYDPTQRTITLDEALQHPFFDLKKK 445
DB 421 EYDPTQRTITLDEALQHPFFDLKKK 445

RESULT 4
```


Db 61 RKNRHCSHQSRKSHRRKRGRSIEDDEGHLCQSGDVLRYEIVDTLGEAGFKV 120
Qy 139 ECIDHGMGMHVAVKIVKNGRYREARSEIOVLEHLNSTDPNSVFCVQMLEWFDHGH 198
Db 121 ECIDHGMGMHVAVKIVKNGRYREARSEIOVLEHLNSTDPNSVFCVQMLEWFDHGH 180
Qy 199 VCIVFELLGLSTYDFIKENSLFPQIDHIROMAYQICQSNFLHNKLTHTDLKPENILF 258
Db 181 VCIVFELLGLSTYDFIKENSLFPQIDHIROMAYQICQSNFLHNKLTHTDLKPENILF 240
Qy 259 VKSDYVVKYNSKMKRDETLKNTDIKVDFGSAFYDDDEHSHSTLVSTRHYRAPEVILALGW 318
Db 241 VKSDYVVKYNSKMKRDETLKNTDIKVDFGSAFYDDDEHSHSTLVSTRHYRAPEVILALGW 300
Qy 319 SOPCDVWSIGCILLIYYLGLFTVQTHDSKEHLAMMERILGPIQHMIOKTRKRYFHHNQ 378
Db 301 SOPCDVWSIGCILLIYYLGLFTVQTHDSKEHLAMMERILGPIQHMIOKTRKRYFHHNQ 360
Qy 379 LDWDEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLFDLVRRMLEYDPTQRIITLDEALQHPF 438
Db 361 LDWDEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLFDLVRRMLEYDPTQRIITLDEALQHPF 420
Qy 439 FDLKKK 445
Db 421 FDLKKK 427

RESULT 7
US-10-109-854-4
; Sequence 4, Application US/10109854
; Publication No. US20020119548A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000758DIV
; CURRENT APPLICATION NUMBER: US/10/109,854
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-109-854-4

Query Match 95.9%; Score 2312; DB 13; Length 427;
Best Local Similarity 100.0%; Pred. No. 5.8e-170;
Matches 427; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 19 HYLEARSINERDYDRRYVDEYRNDYCEGYVPRHYRDIESGYRIHCKSSVRSRRSPK 78
Db 1 HYLEARSINERDYDRRYVDEYRNDYCEGYVPRHYRDIESGYRIHCKSSVRSRRSPK 60
Qy 79 RKNRHCSHQSRKSHRRKRGRSIEDDEGHLCQSGDVLRYEIVDTLGEAGFKV 138
Db 61 RKNRHCSHQSRKSHRRKRGRSIEDDEGHLCQSGDVLRYEIVDTLGEAGFKV 120
Qy 139 ECIDHGMGMHVAVKIVKNGRYREARSEIOVLEHLNSTDPNSVFCVQMLEWFDHGH 198
Db 121 ECIDHGMGMHVAVKIVKNGRYREARSEIOVLEHLNSTDPNSVFCVQMLEWFDHGH 180
Qy 199 VCIVFELLGLSTYDFIKENSLFPQIDHIROMAYQICQSNFLHNKLTHTDLKPENILF 258
Db 181 VCIVFELLGLSTYDFIKENSLFPQIDHIROMAYQICQSNFLHNKLTHTDLKPENILF 240
Qy 259 VKSDYVVKYNSKMKRDETLKNTDIKVDFGSAFYDDDEHSHSTLVSTRHYRAPEVILALGW 318
Db 241 VKSDYVVKYNSKMKRDETLKNTDIKVDFGSAFYDDDEHSHSTLVSTRHYRAPEVILALGW 300
Qy 319 SOPCDVWSIGCILLIYYLGLFTVQTHDSKEHLAMMERILGPIQHMIOKTRKRYFHHNQ 378
Db 301 SOPCDVWSIGCILLIYYLGLFTVQTHDSKEHLAMMERILGPIQHMIOKTRKRYFHHNQ 360
Qy 379 LDWDEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLFDLVRRMLEYDPTQRIITLDEALQHPF 438
Db 361 LDWDEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLFDLVRRMLEYDPTQRIITLDEALQHPF 420

Db 241 VKSDYVVKYNSKMKRDETLKNTDIKVDFGSAFYDDDEHSHSTLVSTRHYRAPEVILALGW 300
Qy 319 SOPCDVWSIGCILLIYYLGLFTVQTHDSKEHLAMMERILGPIQHMIOKTRKRYFHHNQ 378
Db 301 SOPCDVWSIGCILLIYYLGLFTVQTHDSKEHLAMMERILGPIQHMIOKTRKRYFHHNQ 360
Qy 379 LDWDEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLFDLVRRMLEYDPTQRIITLDEALQHPF 438
Db 361 LDWDEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLFDLVRRMLEYDPTQRIITLDEALQHPF 420
Qy 439 FDLKKK 445
Db 421 FDLKKK 427

RESULT 8
US-10-339-656-4
; Sequence 4, Application US/10339656
; Publication No. US20030134319A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000758DIV
; CURRENT APPLICATION NUMBER: US/10/339,656
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/109,854
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-339-656-4

Query Match 95.9%; Score 2312; DB 14; Length 427;
Best Local Similarity 100.0%; Pred. No. 5.8e-170;
Matches 427; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 19 HYLEARSINERDYDRRYVDEYRNDYCEGYVPRHYRDIESGYRIHCKSSVRSRRSPK 78
Db 1 HYLEARSINERDYDRRYVDEYRNDYCEGYVPRHYRDIESGYRIHCKSSVRSRRSPK 60
Qy 79 RKNRHCSHQSRKSHRRKRGRSIEDDEGHLCQSGDVLRYEIVDTLGEAGFKV 138
Db 61 RKNRHCSHQSRKSHRRKRGRSIEDDEGHLCQSGDVLRYEIVDTLGEAGFKV 120
Qy 139 ECIDHGMGMHVAVKIVKNGRYREARSEIOVLEHLNSTDPNSVFCVQMLEWFDHGH 198
Db 121 ECIDHGMGMHVAVKIVKNGRYREARSEIOVLEHLNSTDPNSVFCVQMLEWFDHGH 180
Qy 199 VCIVFELLGLSTYDFIKENSLFPQIDHIROMAYQICQSNFLHNKLTHTDLKPENILF 258
Db 181 VCIVFELLGLSTYDFIKENSLFPQIDHIROMAYQICQSNFLHNKLTHTDLKPENILF 240
Qy 259 VKSDYVVKYNSKMKRDETLKNTDIKVDFGSAFYDDDEHSHSTLVSTRHYRAPEVILALGW 318
Db 241 VKSDYVVKYNSKMKRDETLKNTDIKVDFGSAFYDDDEHSHSTLVSTRHYRAPEVILALGW 300
Qy 319 SOPCDVWSIGCILLIYYLGLFTVQTHDSKEHLAMMERILGPIQHMIOKTRKRYFHHNQ 378
Db 301 SOPCDVWSIGCILLIYYLGLFTVQTHDSKEHLAMMERILGPIQHMIOKTRKRYFHHNQ 360
Qy 379 LDWDEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLFDLVRRMLEYDPTQRIITLDEALQHPF 438
Db 361 LDWDEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLFDLVRRMLEYDPTQRIITLDEALQHPF 420


```
QY 439 FDLLKKK 445
Db 421 FDLLKKK 427

RESULT 9
US-10-801-671-4
; Sequence 4, Application US/10801671
; Publication No. US20040152123A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO00758DIV-111
; CURRENT APPLICATION NUMBER: US/10/801,671
; PRIOR FILING DATE: 2004-03-17
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-801-671-4

Query Match 95.9%; Score 2312; DB 16; Length 427;
Best Local Similarity 100.0%; Pred. No. 5.8e-170;
Matches 427; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 HYLEARSLNERDYRRVYDEYNDYCEGYVPRHYRDIESGYRHCHSKSSVRSRSPK 78
Db 1 HYLEARSLNERDYRRVYDEYNDYCEGYVPRHYRDIESGYRHCHSKSSVRSRSPK 60

QY 79 RKNRHCSSHQSRKSHRRKRSRSIEDDEGHLICQSGDVLRLRYEIVDTLGEAGFGKV 138
Db 61 RKNRHCSSHQSRKSHRRKRSRSIEDDEGHLICQSGDVLRLRYEIVDTLGEAGFGKV 120

QY 139 ECIDHGMGMVAVKIVKNGVRYEARSEIQVLEHNLSTDPNSVFCVQMLEWFDHGH 198
Db 121 ECIDHGMGMVAVKIVKNGVRYEARSEIQVLEHNLSTDPNSVFCVQMLEWFDHGH 180

QY 199 VCIVFELLGLSTYDFIKENSFLPFQIDHIROMAYQICQINFLHNKLTHTDLKPENILF 258
Db 181 VCIVFELLGLSTYDFIKENSFLPFQIDHIROMAYQICQINFLHNKLTHTDLKPENILF 240

QY 259 VKSDYVVKYNSKMKRDERTLKNITDKVVDPGSATYDDSHHSTLVSTRHYRAPEVILALGW 318
Db 241 VKSDYVVKYNSKMKRDERTLKNITDKVVDPGSATYDDSHHSTLVSTRHYRAPEVILALGW 300

QY 319 SQPCDVWSIGCILIEYILGFTVFQTHDSKEHLAMMERILGPIQPMIOKTRKRYFHHNQ 378
Db 301 SQPCDVWSIGCILIEYILGFTVFQTHDSKEHLAMMERILGPIQPMIOKTRKRYFHHNQ 360

QY 379 LDWDEHSSAGRYVRRCKPLKEFMLCHDEHEKLPDLVRRMLEYDPTORITLDEALQHPF 438
Db 361 LDWDEHSSAGRYVRRCKPLKEFMLCHDEHEKLPDLVRRMLEYDPTORITLDEALQHPF 420

QY 439 FDLLKKK 445
Db 421 FDLLKKK 427

RESULT 10
US-09-905-999-25
; Sequence 25, Application US/09905999
; Patent No. US2002010671A1
; GENERAL INFORMATION:
; APPLICANT: ULLRICH, Axel
; APPLICANT: NAYLER, Oliver

; TITLE OF INVENTION: CLK PROTEIN KINASES AND RELATED PRODUCTS AND METHODS
; FILE REFERENCE: 038602/0431
; CURRENT APPLICATION NUMBER: US/09/905,999
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 09/127,248
; PRIOR FILING DATE: 1999-07-31
; PRIOR APPLICATION NUMBER: PCI/IB97/00946
; PRIOR FILING DATE: 1997-06-17
; PRIOR APPLICATION NUMBER: US 08/877,150
; PRIOR FILING DATE: 1997-06-17
; PRIOR APPLICATION NUMBER: US 60/034,286
; PRIOR FILING DATE: 1996-12-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 25
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-905-999-25

Query Match 93.8%; Score 2261.5; DB 9; Length 481;
Best Local Similarity 93.5%; Pred. No. 5.2e-166;
Matches 420; Conservative 8; Mismatches 10; Indels 11; Gaps 2;

QY 8 SH-SVEEDTH-----PSHYLEARSLSNERDYRRYVDEYNDYCEGYVPRHYRD 56
Db 33 SHSSTQENRCHCKPHQKFDSDCHYLEARCLNERDYRRYVDEYNDYCEGYVPRHYRD 92

QY 57 IESGYRHCHSKSSVRSRSPKRNHRCHSHQSRKSHRRKRSRSIEDDEGHLICQSG 116
Db 93 VESTYRIHCHSKSSVRSRSPKRNHRCHSHQSRKSHRRKRSRSIEDDEGHLICQSG 152

QY 117 DVLRLRYEIVDTLGEAGFGKVVECTIDHGMGMVAVKIVKNGVRYEARSEIQVLEHLN 176
Db 153 DVLRLRYEIVDTLGEAGFGKVVECTIDHGMGMVAVKIVKNGVRYEARSEIQVLEHLN 212

QY 177 STDNSVFCVQMLEWFDHGHGVCIVFELLGLSTYDFIKENSFLPFQIDHIROMAYQICQ 236
Db 213 STDNSVFCVQMLEWFDHGHGVCIVFELLGLSTYDFIKENSFLPFQIDHIROMAYQICQ 272

QY 237 SINFLHNKLTHTDLKPENILFVKSDYVVKYNSKMKRDERTLKNITDKVVDPGSATYDD 296
Db 273 SINFLHNKLTHTDLKPENILFVKSDYVVKYNSKMKRDERTLKNITDKVVDPGSATYDD 332

QY 297 HHSTLVSTRHYRAPEVILALGWSQPCDVWSIGCILIEYILGFTVFQTHDSKEHLAMMERI 356
Db 333 HHSTLVSTRHYRAPEVILALGWSQPCDVWSIGCILIEYILGFTVFQTHDSKEHLAMMERI 392

QY 357 LGPIQPMIOKTRKRYFHHNQLDWDEHSSAGRYVRRCKPLKEFMLCHDEHEKLPDLV 416
Db 393 LGPIQPMIOKTRKRYFHHNQLDWDEHSSAGRYVRRCKPLKEFMLCHDEHEKLPDLV 452

QY 417 RRMLEYDPTORITLDEALQHPFOLLKKK 445
Db 453 RRMLEYDPTORITLDEALQHPFOLLKKK 481

RESULT 11
US-10-267-502-355
; Sequence 355, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 355
; LENGTH: 481
; TYPE: PRT
```

2

Db	361	HDRLDWDDEHSSAGRYVSRACKPLKEFMLSQDVEHERLFDLLQKMLEYDPAKRITLREALK	420
Qy	436	HPFFDLLKK	444
Db	421	HPFFDLLKK	429
RESULT 14			
US-10-109-854-5			
; Sequence 5, Application US/10109854			
; Publication No. US20020119548A1			
; GENERAL INFORMATION:			
; APPLICANT: YAN, Chunhua et al.			
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC			
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES			
; TITLE OF INVENTION: THEREOF			
; FILE REFERENCE: CL000758DIV			
; CURRENT APPLICATION NUMBER: US/10/109,854			
; CURRENT FILING DATE: 2002-04-01			
; PRIOR APPLICATION NUMBER: 60/227,470			
; PRIOR FILING DATE: 2000-08-24			
; PRIOR APPLICATION NUMBER: 09/810,671			
; PRIOR FILING DATE: 2001-03-19			
; NUMBER OF SEQ ID NOS: 5			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 5			
; LENGTH: 429			
; TYPE: PRT			
; ORGANISM: Homo sapien			
US-10-109-854-5			
Query Match 78.3%; Score 1887; DB 13; Length 429;			
Best Local Similarity 82.1%; Pred. No. 3.4e-137;			
Matches 352; Conservative 30; Mismatches 45; Indels 2; Gaps 2			
Qy	18	SHYLEARSLNERDYRDRYDEYRNDYCEGVVPRHYHRDIESGYRIHCKSKSVSRSSP	77
Db	1	SHYLESRSINEKDYHSRRYIDYRNDYTGQCEPGHRQRDHESRYQNHSKSGRSY	60
Qy	78	KRK-RNRHCSSH-QSRKSKSRKRSRSTEDDEBGLHCQSGDVLRLRYEIVDTLCEGAFG	135
Db	61	KSKHRIHSTSHRSHGSKSHRKRTRSVDEDBEGLHCQSGDVLRLRYEIVDTLCEGAFG	120
Qy	136	KVVECIDHGMGDMHVAKIVKXNGRYREAAARSEIQVLEHLNSTDPNSVRCVQMLEWFDH	195
Db	121	KVVECIDHKAGRRHAVKIVKXNDRYCEAAARSEIQVLEHLNTDPNSTFRCVQMLEWFEH	180
Qy	196	HGHVCIIVPELLGLSTYDFIKENSFLPFOIDHIROMAYOICOSINPLHKNLTHTDLKPEN	255
Db	181	HGHICIVPELLGLSTYDFIKENGELFFPLDHIRKWAYOICKSVNPLHKNLTHTDLKPEN	240
Qy	256	ILFVKSQYVVKYNSMKRDERETLKNNTD1KVVDFGSATYDDDEHSHSLVSTRHYRAPEVILA	315
Db	241	ILFVQSDYTEAYNPKIKEDERTLINPDIKVDFGSATYDDDEHSHSLVSTRHYRAPEVILA	300
Qy	316	LGSQPCDQWSIGCILIEYILGFTVFPQTHDSKEHLAMMERILGPIQPMIOKTRKRYFHF	375
Db	301	LGSQPCDQWSIGCILIEYILGFTVFPQTHDSKEHLAMMERILGPIQPMIOKTRKRYFHF	360
Qy	376	HNQLDWDDEHSSAGRYVRRRCRKLPEFMLECHDBEHEKLPDLVRRMLEYDPTORITLDEALQ	435
Db	361	HDRLDWDDEHSSAGRYVSRACKPLKEFMLSQDVEHERLFDLLQKMLEYDPAKRITLREALK	420
Qy	436	HPFFDLLKK	444
Db	421	HPFFDLLKK	429

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2005, 23:51:06 ; Search time 170 Seconds
(without alignments)
1340.443 Million cell updates/sec

Title: US-10-801-671-2
Perfect score: 2410
Sequence: 1 MCIPLEASHSVEDTHPSHY.....QRITLDELQHPFDLLKKK 445

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 03.*
1: uniprot_prot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	2315.5	96.1	481 1	CLK4_HUMAN
2	2261.5	93.8	481 1	CLK4_MOUSE
3	1893	78.5	484 2	Q8N5V8
4	1897	78.3	484 1	CLK1_HUMAN
5	1879	78.0	484 2	Q9BGQ1
6	1832	76.0	483 1	CLK1_MOUSE
7	1803	74.8	453 2	Q9NRK6
8	1608	66.7	301 2	Q6AYK7
9	1437	59.6	498 2	Q91YR2
10	1427.5	59.2	499 1	CLK2_MOUSE
11	1417	58.8	499 1	CLK2_HUMAN
12	1398	58.0	491 2	Q7ZV59
13	1321	54.8	490 1	CLK3_HUMAN
14	1317	54.6	490 1	CLK3_RAT
15	1316	54.6	490 2	Q6IRK2
16	1310	54.4	490 1	CLK3_MOUSE
17	1305	54.1	490 2	Q8C1V1
18	1252	52.0	338 2	Q9BRG8
19	1180	49.0	832 1	DOA_DROME
20	1180	49.0	2063 3	Q8GB73
21	1122	46.6	393 2	Q7PN19
22	1119	46.4	376 2	Q7Q4W1
23	1011	42.0	929 2	Q7PQG8
24	969.5	40.2	887 2	Q1917
25	963	40.0	409 2	Q8MLY2
26	884.5	36.7	260 2	Q7PH65
27	848.5	35.2	427 1	AFC2_ARATH
28	839	34.8	431 2	O49967
29	830	34.4	467 1	APC1_ARATH
30	802	33.3	434 2	Q943M7
31	790	32.8	400 1	AFC3_ARATH

32	789	32.7	437	2	Q9M598
33	767	31.8	575	1	LKH1_SCHPO
34	718.5	29.8	737	1	KNS1_YEAST
35	717.5	29.8	658	2	Q6CD14
36	696	28.9	725	2	Q6FVQ2
37	690	28.6	640	2	Q6BU28
38	690	28.6	699	2	Q6CU13
39	673.5	27.9	772	2	Q7RZY1
40	672.5	27.9	673	2	Q7SAG1
41	648.5	26.9	881	2	Q81L19
42	648.5	26.9	881	2	Q9BKN8
43	631.5	26.2	733	2	Q7B9I8
44	598	24.8	1300	2	Q8XOV5
45	550.5	22.8	538	2	Q8BM34

ALIGNMENTS

RESULT 1	CLK4_HUMAN	STANDARD;	PRT;	481 AA.
ID	CLK4_HUMAN			
AC	Q9HAZ1			
DT	10-OCT-2003 (Rel. 42, Created)			
DT	10-OCT-2003 (Rel. 42, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Dual specificity protein kinase CLK4 (EC 2.7.1.37) (EC 2.7.1.112) (CDC like kinase 4).			
GN	Name=CLK4;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AUTOPHOSPHORYLATION, AND MUTAGENESIS OF LYS-189.			
RC	TISSUE=Kidney;			
RX	MEDLINE=21100912; PubMed=11170754; DOI=10.1106/geno.2000.6447;			
RA	Schultz J., Jones T., Bork P., Sheer D., Blencke S., Steyrer S., Wellbrock U., Bevec D., Ullrich A., Wallasch C.			
RT	"Molecular characterization of a cDNA encoding functional human CLK4 kinase and localization to chromosome 5q35."			
RL	Genomics 71:368-370(2001).			
RN	[2]			
RP	ERRATUM.			
RA	Schultz J., Jones T., Bork P., Sheer D., Blencke S., Steyrer S., Wellbrock U., Bevec D., Ullrich A., Wallasch C.			
RL	Genomics 74:251-251(2001).			
RN	[3]			
RP	INTERACTION WITH UBL5.			
RX	MEDLINE=22709234; PubMed=12824502; DOI=10.1110/ps.0382803;			
RA	McNally T., Huang Q., Janis R.S., Liu Z., Olejniczak E.T., Reilly R.M.,			
RT	"Structural analysis of UBL5, a novel ubiquitin-like modifier."			
RL	Protein Sci. 12:1562-1566(2003).			
RN	[4]			
RP	INTERACTION WITH UBL5.			
RX	MEDLINE=22592248; PubMed=12705895; DOI=10.1016/S0006-291X(03)00549-7;			
RA	Kantham L., Kerr-Bayles L., Godde N., Quick M., Webb R., Sunderland T., Bond J., Walder K., Augert G., Collier G.,			
RT	"Beacon interacts with cdc2/cdc28-like kinases."			
RL	Biochem. Biophys. Res. Commun. 304:125-129(2003).			
CC	!- FUNCTION: Phosphorylates serine- and arginine-rich (SR) proteins of the spliceosomal complex may be a constituent of a network of regulatory mechanisms that enable SR proteins to control RNA splicing. Phosphorylates serines, threonines and tyrosines.			
CC	!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.			
CC	!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.			
CC	!- SUBUNIT: Interacts with UBL5.			
CC	!- SUBCELLULAR LOCATION: Nuclear (By similarity).			
CC	!- TISSUE SPECIFICITY: Expressed in liver, kidney, heart, muscle and			

```

CC brain.
CC -|- PTM: Autophosphorylates on all three types of residues.
CC -|- SIMILARITY: Belongs to the Ser/Thr protein kinase family. Lammer
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF294429; AAG10074.1; -
CC HSSP; Q00534; 1B18.
CC Genew; HGNC:13659; CLK4.
CC MIM; 607969; -
CC InterPro; IPR011009; Kinase like.
CC InterPro; IPR000719; Prot kinase.
CC InterPro; IPR008271; Ser Thr pkin.AS.
CC InterPro; IPR002290; Ser_thr_kinase.
CC Pfam; PF00069; Pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TK; 1.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS50011; PROTEIN KINASE DOM; 1.
CC PROSITE; PS00108; PROTEIN KINASE_ST; 1.
CC ATP-binding; Nuclear protein; Phosphorylation; Tyrosine-protein kinase.
CC Serine/threonine-protein kinase; Transferase; (By similarity).
CC FT DOMAIN 159 475
CC FT NP BIND 165 173 ATP (By similarity).
CC FT ACT SITE 286 286 Proton acceptor (By similarity).
CC FT BINDING 189 189 ATP (By similarity).
CC FT MUTAGEN 189 189 K->R: Loss of function.
CC SQ SEQUENCE 481 AA; 57492 MW; F402C36835CDA306 CRC64;

Query Match 96.1%; Score 2315.5; DB 1; Length 481;
Best Local Similarity 96.2%; Pred. No. 2.3e-135;
Matches 432; Conservative 2; Mismatches 4; Indels 11; Gaps 2;

QY 8 SH-SVEEDTH-----PSHYLEARSINRDYRDYRVDYDEYNDYCEGVVPRHYRD 56
Dd |||:|
33 SHSSTQENRHKPHQKESDCHYLEARSINRDYRDYRVDYDEYNDYCEGVVPRHYRD 92
QY 57 IESGVRIHCKSVRRSSPRKRNHCSHQSRKSHRRKRSIEDDEGHLCQSG 116
Dd |||
93 IESGVRIHCKSVRRSSPRKRNHCSHQSRKSHRRKRSIEDDEGHLCQSG 152
QY 117 DVLRLRYEIVDTLGEAGFGKVVCEIDHGMHVAVKIVKNGVRYREARSEIQVLEHLN 176
Dd |||
153 DVLRLRYEIVDTLGEAGFGKVVCEIDHGMHVAVKIVKNGVRYREARSEIQVLEHLN 212
QY 177 STDPSNVRFCVQMLEWFDHGHVCIIVFELLGLSTYDFIKENSFLPQDIIHQMAYQICQ 236
Dd |||
213 STDPSNVRFCVQMLEWFDHGHVCIIVFELLGLSTYDFIKENSFLPQDIIHQMAYQICQ 272
QY 237 SINFLHNKLTHTDLKPENILFKSDYVVKYNSKMKRDERTLKNNTDIKVDFGSAATYDDE 296
Dd |||
273 SINFLHNKLTHTDLKPENILFKSDYVVKYNSKMKRDERTLKNNTDIKVDFGSAATYDDE 332
QY 297 HHSTLVSTRHYRAPEVILALGWSQPCDVWSIGCIIIEYLGFTVFQTHDSKEHLAMMERI 356
Dd |||
333 HHSTLVSTRHYRAPEVILALGWSQPCDVWSIGCIIIEYLGFTVFQTHDSKEHLAMMERI 392
QY 357 LGPIPOHMTQTKRKRYFHHNOLDWDEHSSAGRYVRRCKPKLKEPMLCHDEHEKLFDLV 416
Dd |||
393 LGPIPOHMTQTKRKRYFHHNOLDWDEHSSAGRYVRRCKPKLKEPMLCHDEHEKLFDLV 452
QY 417 RMLEVDPTQRTLDALQHPFDLLKKK 445
Dd |||
453 RMLEVDPTQRTLDALQHPFDLLKKK 481

```

```

RESULT 2
CLK4_MOUSE
ID CLK4_MOUSE STANDARD; PRT; 481 AA.
AC O35493; O35721; Q8CEU9; Q99LU6;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Dual specificity protein kinase CLK4 (EC 2.7.1.37) (EC 2.7.1.112) (CDC
DE like kinase 4).
GN Name=CLK4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=97433710; PubMed=9307018;
RA Nayler O., Stamm S., Ullrich A.;
RT "Characterization and comparison of four serine- and arginine-rich
RL Biochem. J. 326:693-700(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX STRAIN=C57BL/6J; TISSUE=Head;
MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Oeato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verdano R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyszynski B., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai K., Aizawa K., Arakawa T., Fukuda S.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Havaehizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX STRAIN=Czech II; TISSUE=Breast tumor;
MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan K., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

```



```

DR PIR; S53641; S53641.
DR HSP; Q00534; 1B18.
DR Genew; HGNC:2068; CLK1.
DR H-InvDB; HIX0002728; -.
DR MIM; 601951; -.
DR GO; Q0004715; F-non-membrane spanning protein tyrosine kinase; TAS.
DR GO; Q0004674; F-protein serine/threonine kinase activity; TAS.
DR GO; Q0000074; P-cell proliferation; TAS.
DR GO; Q0000074; P-regulation of cell cycle; TAS.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Alternative splicing; ATP-binding; Nuclear protein; Phosphorylation;
Serine/threonine-protein kinase; Transferase; Tyrosine-protein kinase.
FT DOMAIN 161 477
FT NP_BIND 167 175 ATP (By similarity).
FT BINDING 191 191 ATP (By similarity).
FT ACT_SITE 288 288 Proton acceptor (By similarity).
FT VARSPPLIC 131 136 KSHRRK -> MKLLIL (in isoform Short).
FT VARSPPLIC 137 484 Missing (in isoform Short).
FT VARSPPLIC 137 484 /FTId=VSP_004852.
FT SEQUENCE 484 AA; 57205 MW; 30495B486AD0A6B CRC64;
Query Match 78.3%; Score 1887; DB 1; Length 484;
Best Local Similarity 82.1%; Pred. No. 7.4e-109;
Matches 352; Conservative 30; Mismatches 45; Indels 2; Gaps 2;
QY 18 SHYLEARSINERDRDRYVDEYRNDYCEGVPRPHRYDIESGYRTHCSKSSVRSRSP 77
DB 54 SHYLESRSINEXDYGRRYDIEYRNDYTOGCEPGHRQRDHESYQNHSSKSGRSY 113
QY 78 KRX-RNRHCSSH-QSRKSHRRKRSISIEDDEGHLCIQSGDVLRYEIVDTLGGAGF 135
DB 114 KSKHRIHSTSHRRSHGSKSHRRKTRSVEDDEGHLCIQSGDVLRYEIVDTLGGAGF 173
QY 136 KVECIDHGMGMHVAVKIVNVRVREARSIOVLEHLNSTDPNSVFRVCQMLEWFDH 195
DB 174 KVECIDHAGGRHVAVKIVNVDRCYCAARSIOVLEHLNSTDPNSVFRVCQMLEWFEH 233
QY 196 HGHVCLVFEGLLSTYDFIKENSLFPQDHIROMAYQICQSNFLHNKLTHTDLKPEN 255
DB 234 HGHICVFEGLLSTYDFIKENGLFPFLDHIROMAYQICQSNFLHNKLTHTDLKPEN 293
QY 256 ILFVKSDYVYKNSKMRDERTLKNVDIKVDFGSATYDEHSTLVSTRHYRAPVILA 315
DB 294 ILFVQSDYTYANPKIKRDETLINFDIKVDFGSATYDEHSTLVSTRHYRAPVILA 353
QY 316 LGWSQPCDVMSIGCILIEYLGFTVFTQTHDSKEHLAMMERILGPIQPMIQRKRKYFH 375
DB 354 LGWSQPCDVMSIGCILIEYLGFTVFTQTHDSKEHLAMMERILGPIQPMIQRKRKYFH 413
QY 376 HNOLDWDEHSSAGRYVRRCKPKLFKMLCHDEHEKLPDLVRMLDYDPTQITLDEALQ 435
DB 414 HDRLDWEHSSAGRYVSRACKPKLFKMLSDQVHEHRLFDLIQKMLEYDPAKRTITLREALK 473
QY 436 HPFFDLKK 444
DB 474 HPFFDLKK 482
RESULT 5
Q9BGQ1 PRELIMINARY; PRT; 484 AA.
ID Q9BGQ1
AC Q9BGQ1
DT 01-JUN-2001 (TrEMBLrel. 17, Created)

```

```

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Frontal lobe left;
RA Ceada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB056421; BAB33079.1; -
DR HSP; Q00534; 1B18.
DR GO; Q0005524; P-ATP binding; IEA.
DR GO; Q0004674; P-protein serine/threonine kinase activity; IEA.
DR GO; Q0016740; P-transferase activity; IEA.
DR GO; Q0006468; P-protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW ATP-binding; Hypothetical protein; Kinase; Transferase.
SQ SEQUENCE 484 AA; 57318 MW; F5324D4007119878 CRC64;
Query Match 78.0%; Score 1879; DB 2; Length 484;
Best Local Similarity 79.3%; Pred. No. 2.3e-108;
Matches 356; Conservative 30; Mismatches 51; Indels 12; Gaps 4;
QY 8 SHSVEED-----TH-----PSHYLEARSINERDRDRYVDEYRNDYCEGVPRPHRYDI 57
DB 34 SHSSARENRCKYTHSKMCDSHYLESRCINEXDYSRRYDIEYRNDYNOCEPGHRRDH 93
QY 58 ESGYRTHCSKSSVRSRSPKPK-RNRHCSSH-QSRKSHRRKRSISIEDDEGHLCIQS 115
DB 94 ESRYQNHSSKSGRSGRSYKSKHRIHSTSHRRSHGSKSHRRKTRSVEDDEGHLCIQS 153
QY 116 GDLVRLRYEIVDTLGGAGFQVVECIDHGMGMHVAVKIVNVRVREARSIOVLEHL 175
DB 154 GDLVSARYEIVDALGEGAFQVVECIDHAGGRHVAVKIVNVDRCYCAARSIOVLEHL 213
QY 176 NSTDPNSVFRVCQMLEWFDHGHVCIIVFELLGLSTYDFIKENSLFPQDHIROMAYQIC 235
DB 214 NTTDPNSTFRVCQMLEWFEHGHICIVFELLGLSTYDFIKENGLFPFLDHIROMAYQIC 273
QY 236 QSNFLHNKLTHTDLKPENILFVKSDYVYKNSKMRDERTLKNVDIKVDFGSATYDD 295
DB 274 KSNFLHNSKLTHTDLKPENILFVQSDYTYANPKIKRDETLINFDIKVDFGSATYDD 333
QY 296 EHHSTLVSTRHYRAPVILALGWSQPCDVMSIGCILIEYLGFTVFTQTHDSKEHLAMMER 355
DB 334 EHHSTLVSTRHYRAPVILALGWSQPCDVMSIGCILIEYLGFTVFTQTHDSKEHLAMMER 393
QY 356 ILGPIQPMIQRKRKYFHFNQDWEHSSAGRYVRRCKPKLFKMLCHDEHEKLPDL 415
DB 394 ILGPIQPMIQRKRKYFHHDRLDWEHSSAGRYVRRCKPKLFKMLSDQVHEHRLFDL 453
QY 416 VRMLDYDPTQITLDEALQHPFFDLKK 444
DB 454 IQKMLEYDPAKRTITLKEALKHPFFDLKK 482
RESULT 6
CLK1_MOUSE STANDARD; PRT; 483 AA.
ID CLK1_MOUSE

```


464 RTILKEALKHPFFYPLKK 481

Db

RESULT 7

Q9NRL6 PRELIMINARY; PRT; 453 AA.

AC Q9NRL6
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CLK4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RA Zhao M., Song H., Li N., Peng Y., Han Z., Chen Z.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF212224; AAF87326.1; -;
DR PIR; A38643; A38643.
DR HSSP; Q00534; 1B18.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0004688; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0004688; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002230; Ser_Thr_kinase.
DR InterPro; IPR008271; Ser_Thr_kinase.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW ATP-binding; Kinase; Transferase.
SQ SEQUENCE 453 AA; 53806 MW; 3B2AC6A3FFAE9461 CRC64;

Query Match 74.8%; Score 1803; DB 2; Length 453;

Best Local Similarity 80.4%; Pred. No. 1.1e-103; Indels 28; Gaps 3;

Matches 344; Conservative 16; Mismatches 40; Indels 28; Gaps 3;

QY 18 SHYLEARSLNERDYRRDYVDEYRNDYCEGYVPRHYRDIQSYRTHCSKSSVRSRRSP 77

Db 54 SHYLESRSINEKDYHSRRYIDEYRNDYTCGCPGHRQDRHESRYQNHSSKSGRSGSSV 113

QY 78 KRKNRHCSSHQSRKSHRRKRSRSIEDDEEGHLICQSGDVLRLRYEIVDTLGEAGFGKV 137

Db 114 KSKHRIH-----HSTSHRRSH-----GD-----EIVDTLGEAGFGKV 145

QY 138 VECIDHGMGMHVAIVKNGVRYREARSEIQVLEHLNSTDPNSVFCVQMLEWFDHGG 197

Db 146 VECIDHAGRHVAIVKNGVRYREARSEIQVLEHLNSTDPNSVFCVQMLEWFEHGG 205

QY 198 HVCIVFELGLSTYDFIKENSFLPFQIDHIRMAYQICQSFNHLHKNLTHDTLKPENIL 257

Db 206 HICIVFELGLSTYDFIKENGFLPFRLDHRKWAYQICKSVNHLHKNLTHDTLKPENIL 265

QY 258 FVKSQDVVKNKSKKRDERTLKNTDIKVDVFGSATVDDEHSHSTLSTRHYRAPEVILALG 317

Db 266 FVQSDYTEAYNPRIKRDERTLINPDIKVDVFGSATVDDEHSHSTLSTRHYRAPEVILALG 325

QY 318 WSPQCDVMSGTCILIEYVLGFTVFQTHDSKEHLAMMERIIGPIQPMIOKTRCKYFHN 377

Db 326 WSPQCDVMSIGGLNIEYVLGFTVFQTHDSKEHLAMMERIIGPIQPMIOKTRCKYFHN 385

QY 378 QLDWDHSHSAGRYVRRCKPLKEFMLECHDEHEKLFDLVRMLEYDPTQRTIDLEALQHP 437

Db 386 QLDWDHSHSAGRYVRRCKPLKEFMLECHDEHEKLFDLVRMLEYDPTQRTIDLEALQHP 445

QY 438 FFDLLKKK 445

446 FFDLLKKK 453

Db

RESULT 8

Q6AYK7 PRELIMINARY; PRT; 301 AA.

AC Q6AYK7
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RT Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Director MGC Project;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC079006; AAH79006.1; -;
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004688; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002230; Ser_Thr_kinase.
DR InterPro; IPR008271; Ser_Thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; Tyr_KC; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 301 AA; 35671 MW; 7F0B27F043D976C1 CRC64;

Query Match 66.7%; Score 1608; DB 2; Length 301;

Best Local Similarity 98.7%; Pred. No. 8e-92; Indels 0; Gaps 0;

Matches 237; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 145 MDGMHVAIVKNGVRYREARSEIQVLEHLNSTDPNSVFCVQMLEWFDHGHVCIIVE 204

Db 1 MDGLHVAIVKNGVRYREARSEIQVLEHLNSTDPNSVFCVQMLEWFDHGHVCIIVE 60

QY 205 LLGLSTYDFIKENSFLPFQIDHIRMAYQICQSFNHLHKNLTHDTLKPENILFVKSQDV 264

Db 61 LLGLSTYDFIKENSFLPFQIDHIRMAYQICQSFNHLHKNLTHDTLKPENILFVKSQDV 120

290 DLKPENILFVNSDELVTYNLEKKDERSVKSTAVRVVDFSGATFDEHHSTIVSTRHYRA 349

310 PEVILALGWSQPCDWSIGCIIIEYILGFTVFOTHDSKEHLAMWERILGPIPOHMIQKTR 369

350 PEVILELGSQPCDWSIGCIIIEFYVVGFTLFQHDNREHLAMWERILGPIFSMRMKTR 409

370 KKXYFHHNQLDWDHSSAGRYVRRRCPLKEFMLCHDEBEHEKULFDLVRMLLEYDPTQRI 429

410 KOKYFYRGELDWDENTSGRYVRENCKPLRYLTSAEHHQLFDLIESMLMEYEPAKRLT 469

430 IDEALQHPPFDLLK 443

470 LGEALQHPPFARLR 483

RESULT 12

Q7ZY59 PRELIMINARY; PRT; 491 AA.

ID Q7ZY59 AC

DT 01-JUN-2003 (TREMBLrel. 24, Created)

DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)

DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)

DE Clk2-prov protein.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;

OC Xenopodinae; Xenopus.

OX NCBI_TaxID=8355;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;

RX MEDLINE=223188257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Jordan B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Zeehan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Ustin T.B., Tohiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting J., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RC "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences."

RRL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [2]

RP SEQUENCE FROM N.A.

RN TISSUE=Embryo;

RC TISSUE=Embryo;

RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;

RA Klein S.S., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,

RA Richardson P.;

RC "Genetic and genomic tools for Xenopus research: The NIH Xenopus

RT initiative."

RRL Dev. Dyn. 225:384-391(2002).

RN [3]

RP SEQUENCE FROM N.A.

RN TISSUE=Embryo;

RC TISSUE=Embryo;

RX Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

RA EMBL: BC043963; AAH43963.1; -

RA HSP: Q03656; IHW.

DR GO: G0:000524; F:ATP binding; IEA.

DR GO: G0:0004674; F:protein serine/threonine kinase activity; IEA.

DR GO: G0:0016740; F:transferase activity; IEA.

DR GO: G0:0006480; F:protein amino acid phosphorylation; IEA.

DR InterPro: IPRO11009; Kinase like.

DR InterPro: IPR000719; Prot kinase.
 DR InterPro: IPR002290; Ser Thr kinase.
 DR InterPro: IPR008271; Ser Thr pkin_AS.
 DR Pfam: PF00069; Pkinase; 1.
 DR ProDom: PD000001; Prot kinase; 1.
 DR SMART: SM00220; S_TKc_1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
 KW ATP-binding; Kinase; Transferase.
 SQ SEQUENCE 491 AA; 59994 MW; 20870CC59962303P CRC64;
 Query Match 58.0%; Score 1398; DB 2; Length 491;
 Best Local Similarity 59.5%; Pred. No. 1.3e-78;
 Matches 262; Conservative 68; Mismatches 94; Indels 16; Gaps 6;
 QY 6 EASHSVVEEDTHPSHYLEARNLNERDYRDYVDEV-RNDYCEGVPRVHYRDIESGYRTH 64
 DB 50 ERSRSVEE-----RSSDRRAY-DRKYCDSYRRNDYSDRGVYYETDYYD--YK-H 95
 QY 65 CSKSSVRSRRSSPKRKRNRHCSHQSRKSHR-RKRSRSTEDDEGHLICQSGDVLRYR 123
 DB 96 SRDSDSYRSTRKQKRNRRTSRYSQSSRSRQSSRAKSVEDDVVEGLIYHSGDWLQERY 155
 QY 124 EIVDTLGEAGFKVCEIDHGMGMVAVKIVKVGRYREARSEIQVLEHLNSTDPNSV 183
 DB 156 EIVSTLGECTGFRVQCKDHRGGRSVALKIKNVEKYKEARLEINVLEKINEKDPENK 215
 QY 184 FRCVQMLEFDDHGHVCIVPELLGLSTDFIKENSLPQIDHIROMAVQICOSINFLHH 243
 DB 216 HLCVQMFDFDVGHHMCISPELLGLSTDFLKENNYFPPIQVRRMALQOLCOAMKFLHD 275
 QY 244 NKLTHPTDLKPNILFVKSDYVVKYNSKMRDERTLKNTOIKVDFGSAFYDDEHSTLVS 303
 DB 276 NKLTHDLAPENILFVSSDYELRYNNEKKRDERCVKSTDIRVDFGSAFDEHHSTIVS 335
 QY 304 TRHYRAPEVILALGWSQPCDWSIGCILIEYLGLFTVFOTHDSKEHLAMMERILGPIPOH 363
 DB 336 TRHYRAPEVLLELGNQPCDWSVGCIIPEYVYVGTFLFQTHDNREHLAMMERILGPIPSR 395
 QY 364 MIQKTRKRYVHHNQLDWDHESAGRYVRRCKPLKEFVLMCHDEEHEKLFDLVRRMLEYD 423
 DB 396 MIKTRKRYFYHGRUDWDNTSAGRYVRENCKPLRYMMETEEHHQFNFLIEGLLGYE 455
 QY 424 PTQIRITLDEALQHPFFDLK 443
 DB 456 PSKEWTLAALAKHPFNPLK 475
 [1]
 RESULT 13
 CLK3 HUMAN STANDARD; PRT; 490 AA.
 AC P49761; Q9BRS3; Q9BU7;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Dual specificity protein kinase CLK3 (EC 2.7.1.37) (EC 2.7.1.112) (CDC like kinase 3).
 GN Name=CLK3;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RX MEDLINE=95082033; PubMed=7990150;
 RA Hanes J.J., der Kammer H., Klaidiny J.J., Scheit K.H.;
 RT "Characterization by cDNA cloning of two new human protein kinases.
 RT Evidence by sequence comparison of a new family of mammalian protein kinases";
 RL J. Mol. Biol. 244:665-672(1994).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).

RC TISSUE=Cervix, and Placenta;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.C., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SPLICE ISIFORM(S) THAT ARE POTENTIAL NMD TARGET(S).
 RX PubMed=14759258; DOI=10.1186/gb-2004-5-2-r8;
 RA Hillman R.T., Green R.E., Brenner S.E.;
 RT "An unappreciated role for RNA surveillance.";
 RL Genome Biol. 5:RESEARCH008.1-RESEARCH008.16(2004).
 CC -!- FUNCTION: Phosphorylates serine- and arginine-rich (SR) proteins
 CC of the spliceosomal complex may be a constituent of a network of
 CC regulatory mechanisms that enable SR proteins to control RNA
 CC splicing. Phosphorylates serines, threonines and tyrosines (By
 CC similarity).
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1; Synonyms=Long;
 CC IsoId=P49761-1; Sequence=Displayed;
 CC Name=2; Synonyms=Short;
 CC IsoId=P49761-2; Sequence=VSP_004858, VSP_004859;
 CC Notes=Lacks the kinase domain. May be produced at very low levels
 CC due to a premature stop codon in the mRNA, leading to
 CC nonsense-mediated mRNA decay;
 CC Name=3;
 CC IsoId=P49761-3; Sequence=VSP_004860;
 CC -!- PTM: Autophosphorylates on all three types of residues (By
 CC similarity).
 CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. Lammer
 CC subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L29217; AAA61484.1; -;
 DR EMBL; L29220; AAA61483.1; -;
 DR EMBL; BC002555; AAH02555.1; -;
 DR EMBL; BC006103; AAH06103.1; -;
 DR EMBL; BC019881; AAH19881.1; -;
 DR EMBL; S53639; S53639.
 DR PIR; S53640; S53640.
 DR HSSP; Q00534; 1B18.
 DR Genew; HGNC:2071; CLK3.
 DR H-InvDB; HIX0018051; -;
 DR MIM; 602990; -;
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; TAS.
 DR


```

Db 199 RLEINVKKIKKEDKENKFLCVLMSDFNFHGHMCIAPELLGKNTPEFLKENNFQYPLP 258
QY 226 HIROMAYQICQSFNFHHNKLHTDLKPNILFVKSDYVVKVNSKMKRDLTKNTDIKV 285
Db 259 HVHRMAYQLCHALRFLHNLQHTDLKPNILFVNSEFTLYNEHKSCEKSVKNTSIRV 318
QY 286 VFGSATYDDEHSTLVSTRHVRAPDEVILALGWSQPCDVWSIGCILLEYLFTVPTQTHD 345
Db 319 ADFGSATFDHEHHTTIVATRHVRPPEVILELGLWAQPCDVWSIGCILLEYLFTVPTQTHE 378
QY 346 SKEHLAMMERILGPIPOHMIQKTRKRYFHHNOLDWDEHSSAGRYVRRCKPLKEFMLCH 405
Db 379 NREHLVWMEKILGPIPSHMHTRKQKRYFGLVWDEHSSDGRYVYKENCCKPLKSLMLQD 438
QY 406 DEHEKFLDVLRRMLEYDPTQRTILDEALQHPFF 439
Db 439 SLEHVQLFDLRRMLEFDPQSRTITLAEALLHPFF 472

RESULT 15
Q6IRK2
ID Q6IRK2 PRELIMINARY; PRT; 490 AA.
AC Q6IRK2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Clk3 protein.
GN Names=Clk3;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelestomii;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalusz D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC070891; AAH70891.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004574; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser Thr kinase.
DR InterPro; IPR008271; Ser Thr_kin_A5.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; Pkinase; 1.

```

```

DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW ATP-binding; Kinase; Transferase.
SQ SEQUENCE 490 AA; 58485 MW; 6B10CA96A64AB19A CRC64;

Query Match 54.6%; Score 1316; DB 2; Length 490;
Best Local Similarity 55.1%; Pred. No. 1.6e-73;
Matches 250; Conservative 60; Mismatches 102; Indels 42; Gaps 6;

QY 4 PLEASHSVEEDTHPSHYLEARSINERDYDRRYVDEYRNDYCEGYVP-----R 51
Db 43 PPRSRSRSHDRIP-----YORRYREHRDSOTYR---CEERSPFGEDCYGSSRSR 90
QY 52 HYHRDIESG-YRI-----HCKSSSVRSRSPKRNHCHSSHQSSKSHRRKRSRSIED 105
Db 91 HRRSRERGPYTRKHAHHCHK-----RRTSCSSASSRSQSSKRSRSVED 138
QY 106 DEEGHLICQSGDVLARARYEIVDTLGEAGFKVVECIDHGMGMHVAVKIVKNGVYREA 165
Db 139 DKEGHLVCRIGDWLQERYEIVGNLGEETGKVECLDHARGKSQVALKIIRVGVKREA 198
QY 166 RSEIQVLEHLNSTDPNSVRCVOMLEWFDHGHVCIIVFELLGLSTVDFIKENSFLPQID 225
Db 199 RLEINVKKIKKEDKENKFLCVLMSDFNFHGHMCIAPELLGKNTPEFLKENNFQYPLP 258
QY 226 HIROMAYQICQSFNFHHNKLHTDLKPNILFVKSDYVVKVNSKMKRDLTKNTDIKV 285
Db 259 HVHRMAYQLCHALRFLHNLQHTDLKPNILFVNSEFTLYNEHKSCEKSVKNTSIRV 318
QY 286 VFGSATYDDEHSTLVSTRHVRAPDEVILALGWSQPCDVWSIGCILLEYLFTVPTQTHD 345
Db 319 ADFGSATFDHEHHTTIVATRHVRPPEVILELGLWAQPCDVWSIGCILLEYLFTVPTQTHE 378
QY 346 SKEHLAMMERILGPIPOHMIQKTRKRYFHHNOLDWDEHSSAGRYVRRCKPLKEFMLCH 405
Db 379 NREHLVWMEKILGPIPSHMHTRKQKRYFGLVWDEHSSDGRYVYKENCCKPLKSLMLQD 438
QY 406 DEHEKFLDVLRRMLEYDPTQRTILDEALQHPFF 439
Db 439 SLEHVQLFDLRRMLEFDPQSRTITLAEALLHPFF 472

```

Search completed: March 13, 2005, 00:08:00

Job time : 173 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2005, 04:53:38 ; Search time 10125.3 Seconds
(without alignments)
11265.245 Million cell updates/sec

Title: US-10-801-671-1

Perfect score: 2354
Sequence: I gccagctgggttacttta.....taaaaaaacaaaaaaac 2354

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.on.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2354	100.0	2354	6 AR232168	AR232168 Sequence
2	2354	100.0	2354	6 AR406007	AR406007 Sequence
3	2354	100.0	2354	6 AR534243	AR534243 Sequence
4	2354	100.0	2354	6 AX392827	AX392827 Sequence
5	2052.6	87.2	2446	6 AR077239	AR077239 Sequence
6	2052.6	87.2	2446	6 BD107903	BD107903 Protein k
7	2048	87.0	2488	6 AX056410	AX056410 Sequence
8	1632	69.3	1814	9 AF294429	AF294429 Homo sapi
9	1629.4	69.2	1792	6 AX961896	AX961896 Sequence
10	1318.4	56.0	3492	9 BC063116	BC063116 Homo sapi
11	1284	54.5	1446	12 AY335726	AY335726 Synthetic
12	1277.6	54.3	1914	10 BC012675	BC012675 Mus muscu
13	1277.4	54.3	1881	6 AX961894	AX961894 Sequence
14	1185	50.3	1222	6 BD260616	BD260616 49 human
15	1178.2	50.1	1456	6 AR380154	AR380154 Sequence
16	1118.8	47.5	1865	6 AX961898	AX961898 Sequence
17	1097.4	46.6	1549	10 AF033566	AF033566 Mus muscu
18	1097	46.6	21234	6 AR232169	AR232169 Sequence
19	1097	46.6	21234	6 AR406008	AR406008 Sequence

20	1097	46.6	21234	6 AR534244	AR534244 Sequence
21	1097	46.6	21234	6 AX392829	AX392829 Sequence
c 22	1097	46.6	159266	2 AC026097	AC026097 Homo sapi
c 23	1097	46.6	170624	9 AC022096	AC022096 Homo sapi
24	1009.6	42.9	1445	10 BC002220	BC002220 Mus muscu
25	995	42.3	1150	6 CQ718110	CQ718110 Sequence
26	956.8	40.6	1414	10 AY653308	AY653308 Mus muscu
27	910	38.7	1416	10 BC079006	BC079006 Rattus no
28	881.6	37.5	1745	5 BX931976	BX931976 Gallus ga
29	824.4	35.0	1864	9 AB056421	AB056421 Macaca fa
30	814.8	34.6	1757	6 CQ715616	CQ715616 Sequence
31	813.2	34.5	1773	9 BC031549	BC031549 Homo sapi
32	811.6	34.5	1834	9 HUMCLK18	L22219 Homo sapien
33	768.8	32.7	1705	10 NMCLMRNA	X57186 Mouse clk m
34	765.6	32.5	1740	10 MUSSTYKIN	M38381 Mouse serin
35	692	29.4	1946	5 CR848121	CR848121 Xenopus t
36	677.8	28.8	1544	5 BX936001	BX936001 Gallus ga
37	669.4	28.4	1017	5 BX929601	BX929601 Gallus ga
38	644.4	27.4	1743	9 HUMCLK1A	L29222 Homo sapien
39	643.2	27.3	1750	9 HUMKINCD	MS9287 Human prote
40	640.6	27.2	2254	6 AR531674	AR531674 Sequence
c 41	639.4	27.2	1535	11 BV179987	BV179987 sgml10751
42	630.6	26.8	1643	10 MUSMCLK	L29221 Mus musculu
43	628.2	26.7	4862	10 BC062887	BC062887 Mus muscu
44	617	26.2	3056	10 BC081942	BC081942 Rattus no
45	595.2	25.3	1634	5 BX929386	BX929386 Gallus ga

ALIGNMENTS

RESULT 1	AR232168	Sequence 1 from patent US 6455291.	2354 bp	DNA	linear	PAT 20-DEC-2002
LOCUS	AR232168					
DEFINITION	AR232168					
ACCESSION	AR232168.1	GI:27274056				
VERSION						
KEYWORDS						
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 2354)					
AUTHORS	Yan, C., Ye, J., Ketchum, K.A., Di Francesco, V. and Beasley, E.M.					
TITLE	Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins, and uses thereof					
JOURNAL	Patent: US 6455291-A 1 24-SEP-2002;					
FEATURES	Location/Qualifiers					
source	1..2354					
ORIGIN	/organism="unknown"					
	/mol_type="genomic DNA"					

Query Match	100.0%	Score	2354;	DB	6;	Length	2354;
Best Local Similarity	100.0%	Pred. No.	0;				
Matches	2354;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	GCCAGCTGGGGTTACTTTAAAAACATGCTCCATGTGCATCCCTCTTGAAGCTTCGCACT	60				
Db	1	GCCAGCTGGGGTTACTTTAAAAACATGCTCCATGTGCATCCCTCTTGAAGCTTCGCACT	60				
Qy	61	CTGTTGAAGAGGACCACTCATCCCCAGTCATATTTAGAAAGCAAGGTCCTTGAATAGCGAG	120				
Db	61	CTGTTGAAGAGGACCACTCATCCCCAGTCATATTTAGAAAGCAAGGTCCTTGAATAGCGAG	120				
Qy	121	ATTATCGGACCGAGATACGTTTACGATACAGGATGACTACTGTGAAGGATATGTTTC	180				
Db	121	ATTATCGGACCGAGATACGTTTACGATACAGGATGACTACTGTGAAGGATATGTTTC	180				
Qy	181	CTAGACATTATCAGAGACATTCGAAGCGGGTATCGAATCCACTGCAAGTAACTTTCAG	240				
Db	181	CTAGACATTATCAGAGACATTCGAAGCGGGTATCGAATCCACTGCAAGTAACTTTCAG	240				
Qy	241	TCCGACGAGGAGGAGCAGTCTCTAAAAGGAGCGCAATAGACACTGTTTCAAGTCATCAGT	300				

Db 241 TCCGAGCAGGAGAGCAGTCTTAAAGGAAAGCGCAATAGACACTGTTCAGAGTCATCAGT 300
QY |||||
Db 301 CACGTTCCAGAGGCCACCGAAGGAAAGATCCAGAGTATAGAGATGATCAGAGGGTC 360
QY |||||
Db 301 CACGTTCCAGAGGCCACCGAAGGAAAGATCCAGAGTATAGAGATGATCAGAGGGTC 360
QY |||||
Db 361 ACCTGATCTGTCAAGTGGAGACGTTCTTAAGAGCAAGATATGAATCGTGACACTTTGG 420
QY |||||
Db 361 ACCTGATCTGTCAAGTGGAGACGTTCTTAAGAGCAAGATATGAATCGTGACACTTTGG 420
QY |||||
Db 421 GTGAAGGAGCCTTTGGCAAGTTGAGAGTGCATTGATCATGGCATGGATGGCATGCAATG 480
QY |||||
Db 421 GTGAAGGAGCCTTTGGCAAGTTGAGAGTGCATTGATCATGGCATGGATGGCATGCAATG 480
QY |||||
Db 481 TAGCAGTCAAAATCGTAAGAAATGTAGCCGTTACCGTGAAGCAGCTCGTTAGAAATCC 540
QY |||||
Db 481 TAGCAGTCAAAATCGTAAGAAATGTAGCCGTTACCGTGAAGCAGCTCGTTAGAAATCC 540
QY |||||
Db 541 AAGTATTAGAGCAGCTTAAATAGTACTGATCCAAATAGTGTCTCCGATGTCTCCAGATGC 600
QY |||||
Db 541 AAGTATTAGAGCAGCTTAAATAGTACTGATCCAAATAGTGTCTCCGATGTCTCCAGATGC 600
QY |||||
Db 601 TAGAATGGTTTGTGATCATCATGGTCAATGGTTGTAATGCTGTTGAACTACTGGGACTTAGTA 660
QY |||||
Db 601 TAGAATGGTTTGTGATCATCATGGTCAATGGTTGTAATGCTGTTGAACTACTGGGACTTAGTA 660
QY |||||
Db 661 CTTACGATTTCAATTAAGAAACAGCTTTCTGCCATTTCAATTAAGAAATTAACCCATACAG 720
QY |||||
Db 661 CTTACGATTTCAATTAAGAAACAGCTTTCTGCCATTTCAATTAAGAAATTAACCCATACAG 720
QY |||||
Db 721 TGGCGTATCAGATCTGCCAGTCAATAAATTTTTTACATCAATAAATAAATTAACCCATACAG 780
QY |||||
Db 721 TGGCGTATCAGATCTGCCAGTCAATAAATTTTTTACATCAATAAATAAATTAACCCATACAG 780
QY |||||
Db 781 ATCTGAAGCCTGAAATATTTTGTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTA 840
QY |||||
Db 781 ATCTGAAGCCTGAAATATTTTGTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTA 840
QY |||||
Db 841 AAATGAACCGTGAAGACGACCTGAAACACACAGATATCAAAAGTTGTTGACCTTTGGAA 900
QY |||||
Db 841 AAATGAACCGTGAAGACGACCTGAAACACACAGATATCAAAAGTTGTTGACCTTTGGAA 900
QY |||||
Db 901 GTGCAACGTATGATGATCAACATCACAGTACTTTGGTGTCTACCCGGCAGCTACAGAGCTC 960
QY |||||
Db 901 GTGCAACGTATGATGATCAACATCACAGTACTTTGGTGTCTACCCGGCAGCTACAGAGCTC 960
QY |||||
Db 961 CCGAGGTCAATTTGGCTTTAGGTTGGTCTCAGCCCTTGTGATGTTTGGAGCATAGGTTGCA 1020
QY |||||
Db 961 CCGAGGTCAATTTGGCTTTAGGTTGGTCTCAGCCCTTGTGATGTTTGGAGCATAGGTTGCA 1020
QY |||||
Db 1021 TTCTTATTGAATATTACCTTGGTTTCAGTCTTTTCAGATCTCATGATAGTAAGAGCACC 1080
QY |||||
Db 1021 TTCTTATTGAATATTACCTTGGTTTCAGTCTTTTCAGATCTCATGATAGTAAGAGCACC 1080
QY |||||
Db 1081 TGGCAATCATGGAACGAATATTAGAGCCCATACCAACACATGATTCAGAAAAACAAGAA 1140
QY |||||
Db 1081 TGGCAATCATGGAACGAATATTAGAGCCCATACCAACACATGATTCAGAAAAACAAGAA 1140
QY |||||
Db 1141 AACGCAAGTATTTTCAACATTAACAGTATAGTTGGATGAACACAGTTCTGCTGGTAGAT 1200
QY |||||
Db 1141 AACGCAAGTATTTTCAACATTAACAGTATAGTTGGATGAACACAGTTCTGCTGGTAGAT 1200
QY |||||
Db 1201 ATGTTTAGAGAGCGCTGCAACCGTTGAAGGATTTATGCTTTGTCATCATGAAGAACATG 1260
QY |||||
Db 1201 ATGTTTAGAGAGCGCTGCAACCGTTGAAGGATTTATGCTTTGTCATCATGAAGAACATG 1260
QY |||||
Db 1261 AGAAACTGTTTGACCTGTTGCAAGAAATTTAGAAATATGATCCAACTCAAGAAATTTACCT 1320
QY |||||
Db 1261 AGAAACTGTTTGACCTGTTGCAAGAAATTTAGAAATATGATCCAACTCAAGAAATTTACCT 1320
QY |||||
Db 1321 TGGATGAAGCATTTGAGCATCTTTCTTTGACTTATTAAGAAATGAATGGGAATC 1380
QY |||||

RESULT 2
AR406007

Db 1321 TGGATGAAGCATTTGAGCATCTTTCTTTGACTTATTAAGAAATGAATGGGAATC 1380
QY |||||
Db 1381 AGTGGTCTTACTATATCTTCTCTAGAAAGAGATTACTTAAGACTGTGTGCTCAACTAAA 1440
QY |||||
Db 1381 AGTGGTCTTACTATATCTTCTCTAGAAAGAGATTACTTAAGACTGTGTGCTCAACTAAA 1440
QY |||||
Db 1441 CATTTCTAATATTTTGTAAACATTTAAATTTTGTGACAGTTAAGTGAATAATTTGTAATG 1500
QY |||||
Db 1441 CATTTCTAATATTTTGTAAACATTTAAATTTTGTGACAGTTAAGTGAATAATTTGTAATG 1500
QY |||||
Db 1501 TTTTGTATCAATAGCATAAATTAACCTTGTAAAGCAAGTATGCTCTTGATTAAGCAATGAA 1560
QY |||||
Db 1501 TTTTGTATCAATAGCATAAATTAACCTTGTAAAGCAAGTATGCTCTTGATTAAGCAATGAA 1560
QY |||||
Db 1561 AAAATTAATAATTTTCTTTTCTTTTAAATACCACTTTTAAATACCTTTGAAATATCCTT 1620
QY |||||
Db 1561 AAAATTAATAATTTTCTTTTCTTTTAAATTAACCACTTTTAAATACCTTTGAAATATCCTT 1620
QY |||||
Db 1621 TGTGTCAGTGAATAATGTGATTCATCTTGCTTTTGTGACATGAGGTCACCTCTGAAGT 1680
QY |||||
Db 1621 TGTGTCAGTGAATAATGTGATTCATCTTGCTTTTGTGACATGAGGTCACCTCTGAAGT 1680
QY |||||
Db 1681 GAATTTTCTGAGTAAAGGAAATCTTTGACCTTATATTTCTTAAAGGAATATTTCTTTA 1740
QY |||||
Db 1681 GAATTTTCTGAGTAAAGGAAATCTTTGACCTTATATTTCTTAAAGGAATATTTCTTTA 1740
QY |||||
Db 1741 TATACTTTCAAAATTTAGAACTTTAACTTTTAAAGTTTCTTCTGTAATTTGTTGAACGGGTG 1800
QY |||||
Db 1741 TATACTTTCAAAATTTAGAACTTTAACTTTTAAAGTTTCTTCTGTAATTTGTTGAACGGGTG 1800
QY |||||
Db 1801 ATTATTAATTAATCTAGATAAGCAGGTACTAGAAACCAAACTCAGAAAAATGTTTACTGT 1860
QY |||||
Db 1801 ATTATTAATTAATCTAGATAAGCAGGTACTAGAAACCAAACTCAGAAAAATGTTTACTGT 1860
QY |||||
Db 1861 TAGAATTTCTATTAATTTTAAAGTTTGTGATTTCTTTTCAITTTGGGTGATGTCAGGTTGATA 1920
QY |||||
Db 1861 TAGAATTTCTATTAATTTTAAAGTTTGTGATTTCTTTTCAITTTGGGTGATGTCAGGTTGATA 1920
QY |||||
Db 1921 ACCAGACATTCATGGAAGGCGATGAGTTGTCATTTGTCAGTTGCTGTAATTAATAAAACC 1980
QY |||||
Db 1921 ACCAGACATTCATGGAAGGCGATGAGTTGTCATTTGTCAGTTGCTGTAATTAATAAAACC 1980
QY |||||
Db 1981 ACATACACATTTTAAAGATTAATAATCTAACTGGAAGTCAAGCTTGGAAAAATGGACAT 2040
QY |||||
Db 1981 ACATACACATTTTAAAGATTAATAATCTAACTGGAAGTCAAGCTTGGAAAAATGGACAT 2040
QY |||||
Db 2041 TTCCAAAGTATGTTTGGTGGTCAAGATATAAATAAGAAATTTCTGATGAGAGGTTTCAG 2100
QY |||||
Db 2041 TTCCAAAGTATGTTTGGTGGTCAAGATATAAATAAGAAATTTCTGATGAGAGGTTTCAG 2100
QY |||||
Db 2101 TTTTAAATACCAAGTCTTTAGGAGTCTTAAACATTTGGCCAGCATCTGTTTATCAAAATGACA 2160
QY |||||
Db 2101 TTTTAAATACCAAGTCTTTAGGAGTCTTAAACATTTGGCCAGCATCTGTTTATCAAAATGACA 2160
QY |||||
Db 2161 TAAATACGTAACCTATAAGAAATTAAGTTTAAATTAAGGCAATTTTATGCTGTGATTAAT 2220
QY |||||
Db 2161 TAAATACGTAACCTATAAGAAATTAAGTTTAAATTAAGGCAATTTTATGCTGTGATTAAT 2220
QY |||||
Db 2221 TCTTACCGGAGAAAGAGGATTTGATTTGGAAGAGGATTTGGGAAGAAAGTGTGCTGTAAT 2280
QY |||||
Db 2221 TCTTACCGGAGAAAGAGGATTTGATTTGGAAGAGGATTTGGGAAGAAAGTGTGCTGTAAT 2280
QY |||||
Db 2281 TTCCAGAAATTTAAATTTGATTTGATTTACATAAACTTTTGTGACTTTAGAAAAAATAAAAA 2340
QY |||||
Db 2281 TTCCAGAAATTTAAATTTGATTTGATTTACATAAACTTTTGTGACTTTAGAAAAAATAAAAA 2340
QY |||||
Db 2341 AACAAAAAANAAC 2354
QY |||||
Db 2341 AACAAAAAANAAC 2354
QY |||||

LOCUS AR406007 2354 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6630337.
ACCESSION AR406007
VERSION AR406007.1 GI:4015113
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 2354)
AUTHORS Yan, C., Ye, J., Ketchum, K.A., Di Francesco, V. and Beasley, E.M.
TITLE Isolated human kinase proteins, nucleic acid molecules encoding
human kinase proteins, and uses thereof
JOURNAL Patent: US 6630337-A 1 07-OCT-2003;
FEATURES Location/Qualifiers
source 1..2354
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 2354; DB 6; Length 2354;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCAGCTGGGTACTTTAAAAAACATGCTCCATGTGCATCCCTCTTGAAGCTTCGCAC 60
DB 1 GCCAGCTGGGTACTTTAAAAAACATGCTCCATGTGCATCCCTCTTGAAGCTTCGCAC 60
QY 61 CTGTTGAAGAGACACTCATCCAGTCAATTTTAAAGACGAGGTCCTTGAATGAGCGAG 120
DB 61 CTGTTGAAGAGACACTCATCCAGTCAATTTTAAAGACGAGGTCCTTGAATGAGCGAG 120
QY 121 ATTATCGGAGCGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTC 180
DB 121 ATTATCGGAGCGGAGATAGCTTGAAGAAACAGGAATGACTACTGTGAAGGATATGTC 180
QY 181 CTAGACATATACAGAGACATTTAAAGGGGTATCGAATCCACTGCGAGTAAATCTTCAG 240
DB 181 CTAGACATATACAGAGACATTTAAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAG 240
QY 241 TCCGACAGGAGAGAGAGTCTTAAAGAGCGCAATAGACACTGTTCAAGTCAATCAGT 300
DB 241 TCCGACAGGAGAGAGAGTCTTAAAGAGCGCAATAGACACTGTTCAAGTCAATCAGT 300
QY 301 CACGTTTGAAGAGCGCACCGAAGAAAGATCAGGAGTATAGAGGATGATGAGGAGGTC 360
DB 301 CACGTTTGAAGAGCGCACCGAAGAAAGATCAGGAGTATAGAGGATGATGAGGAGGTC 360
QY 361 ACCTGATCTGCAAGTGGAGACGTTCTAAGAGCAAGATATGAATTCGTGGACACTTTGG 420
DB 361 ACCTGATCTGCAAGTGGAGACGTTCTAAGAGCAAGATATGAATTCGTGGACACTTTGG 420
QY 421 GTGAAGGAGCCTTTGGCAAGGTTGTAGAGTGCATTTGATGATGATGATGATGATGATG 480
DB 421 GTGAAGGAGCCTTTGGCAAGGTTGTAGAGTGCATTTGATGATGATGATGATGATGATG 480
QY 481 TAGCAGTGAATAATCGTAAAAATGTAGGCGGTTACCGTGAAGCAGCTCGTTCAGAAATCC 540
DB 481 TAGCAGTGAATAATCGTAAAAATGTAGGCGGTTACCGTGAAGCAGCTCGTTCAGAAATCC 540
QY 541 AAGTATTAGACACTTAAATAGTACTGATCCCAATAGTGTCTTCCGATGATGATGATGATG 600
DB 541 AAGTATTAGACACTTAAATAGTACTGATCCCAATAGTGTCTTCCGATGATGATGATGATG 600
QY 601 TAGAATGTTTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
DB 601 TAGAATGTTTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY 661 CTTACGATTTCAATTAAGAAACAGCTTCTGCCATTTCAAAATGACCAACATCAGGCAGAG 720
DB 661 CTTACGATTTCAATTAAGAAACAGCTTCTGCCATTTCAAAATGACCAACATCAGGCAGAG 720
QY 721 TGGCGTATCAGATCGCCAGTCAATAAATTTTATCATCAATAAATTAATTAATTAATTAAT 780
DB 721 TGGCGTATCAGATCGCCAGTCAATAAATTTTATCATCAATAAATTAATTAATTAATTAAT 780
QY 781 ATCTGAAGCCTGAAATAATTTTGTGTGAAGTCTGACTATGACTCAAAATATAATTTCTA 840
DB 781 ATCTGAAGCCTGAAATAATTTTGTGTGAAGTCTGACTATGACTCAAAATATAATTTCTA 840
QY 841 AAATGAACGTTGATGAACGACACCTGAAACACACAGATATCAAAAGTTGTGACTTTGGA 900
DB 841 AAATGAACGTTGATGAACGACACCTGAAACACACAGATATCAAAAGTTGTGACTTTGGA 900
QY 901 GTGCAACGTTGATGAACGACATCACTGCTTGTGTGTCTACCGGCACTACAGAGCTC 960
DB 901 GTGCAACGTTGATGAACGACATCACTGCTTGTGTGTCTACCGGCACTACAGAGCTC 960
QY 961 CCGAGGTCATTTTGGCTTTAGGTTGGTCTCAGCCCTTGTGATGTTTGGAGCATAGTTGCA 1020
DB 961 CCGAGGTCATTTTGGCTTTAGGTTGGTCTCAGCCCTTGTGATGTTTGGAGCATAGTTGCA 1020
QY 1021 TTCTTATTTGAATATATACCTTGGTTTTCAGCTCTTTCAGACTCATGATAGTAAGAGCACC 1080
DB 1021 TTCTTATTTGAATATATACCTTGGTTTTCAGCTCTTTCAGACTCATGATAGTAAGAGCACC 1080
QY 1081 TGGCAATGATGAACGAAATATTAGGACCCATACCAACACATGATTCAGAAACACAGAA 1140
DB 1081 TGGCAATGATGAACGAAATATTAGGACCCATACCAACACATGATTCAGAAACACAGAA 1140
QY 1141 AACCAAGTATTTTCCACCATTAACAGCTAGATGAGGATGAAACACACAGTTCTGCTGTAG 1200
DB 1141 AACCAAGTATTTTCCACCATTAACAGCTAGATGAGGATGAAACACACAGTTCTGCTGTAG 1200
QY 1201 ATGTTAGAGAGCGTGCACCAACCGTTGAAGAAATTTATGCTTTGTGATGATGATGATG 1260
DB 1201 ATGTTAGAGAGCGTGCACCAACCGTTGAAGAAATTTATGCTTTGTGATGATGATGATG 1260
QY 1261 AGAACTGTTTGAAGTGGTTCGAAAGAAATGTTAGAAATATGATCCCACTCAAGAAATTA 1320
DB 1261 AGAACTGTTTGAAGTGGTTCGAAAGAAATGTTAGAAATATGATCCCACTCAAGAAATTA 1320
QY 1321 TGGATGAAGCAATGCGAGCATCTCTTCTTGTGCTTATTTAAAGAAATGAAATGGAATC 1380
DB 1321 TGGATGAAGCAATGCGAGCATCTCTTCTTGTGCTTATTTAAAGAAATGAAATGGAATC 1380
QY 1381 AGTGTCTTATATATATCTCTAGAAAGATTTCTTAAAGATGCTGTGCTCAACTTAA 1440
DB 1381 AGTGTCTTATATATATCTCTAGAAAGATTTCTTAAAGATGCTGTGCTCAACTTAA 1440
QY 1441 CATCTTAATATTTTGTAAACATTTAAATTTATTTGTACAGTTAAGTGAATATTTGTATG 1500
DB 1441 CATCTTAATATTTTGTAAACATTTAAATTTATTTGTACAGTTAAGTGAATATTTGTATG 1500
QY 1501 TTTTGTATCAATAGCATAAATTAATCTTGTAAAGCAAGTATGCTTGTGATAATGCAATGAA 1560
DB 1501 TTTTGTATCAATAGCATAAATTAATCTTGTGTAAAGCAAGTATGCTTGTGATAATGCAATGAA 1560
QY 1561 AAATTAATAATTTTCTTTTGAATTAACATTTTAAATACCTTTTGAATATCTTT 1620
DB 1561 AAATTAATAATTTTCTTTTGAATTAACATTTTAAATACCTTTTGAATATCTTT 1620
QY 1621 TGTGTCCAGTCAATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680
DB 1621 TGTGTCCAGTCAATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680
QY 1681 GATTTTTTTTTCAGTAAAGGAAATCTTGCATCTTATTTATTTTAAAGGAAATTTCTTTA 1740
DB 1681 GATTTTTTTTTCAGTAAAGGAAATCTTGCATCTTATTTATTTTAAAGGAAATTTCTTTA 1740
QY 1741 TATATCTCAAAATTTAGAACTTAACTTTTAAAGTCTTCTTCTGTAATTTGTTGAACGGGTG 1800
DB 1741 TATATCTCAAAATTTAGAACTTAACTTTTAAAGTCTTCTTCTGTAATTTGTTGAACGGGTG 1800
QY 1801 ATTATTTATTAATCTAGATAGCAGGTAAGAAACCAAACTCAGAAATGTTTACTGT 1860
DB 1801 ATTATTTATTAATCTAGATAGCAGGTAAGAAACCAAACTCAGAAATGTTTACTGT 1860

Db 721 TGGCGTATCAGATCTGCCAGTCAATAAATTTTATCATCAATAAATTAATTAATTAATTAAT 780
QY 781 ATCTGAAGCCTGAAATAATTTTGTGTGAAGTCTGACTATGACTCAAAATATAATTTCTA 840
Db 781 ATCTGAAGCCTGAAATAATTTTGTGTGAAGTCTGACTATGACTCAAAATATAATTTCTA 840
QY 841 AAATGAACGTTGATGAACGACACCTGAAACACACAGATATCAAAAGTTGTGACTTTGGA 900
Db 841 AAATGAACGTTGATGAACGACACCTGAAACACACAGATATCAAAAGTTGTGACTTTGGA 900
QY 901 GTGCAACGTTGATGAACGACATCACTGCTTGTGTGTCTACCGGCACTACAGAGCTC 960
Db 901 GTGCAACGTTGATGAACGACATCACTGCTTGTGTGTCTACCGGCACTACAGAGCTC 960
QY 961 CCGAGGTCATTTTGGCTTTAGGTTGGTCTCAGCCCTTGTGATGTTTGGAGCATAGTTGCA 1020
Db 961 CCGAGGTCATTTTGGCTTTAGGTTGGTCTCAGCCCTTGTGATGTTTGGAGCATAGTTGCA 1020
QY 1021 TTCTTATTTGAATATATACCTTGGTTTTCAGCTCTTTCAGACTCATGATAGTAAGAGCACC 1080
Db 1021 TTCTTATTTGAATATATACCTTGGTTTTCAGCTCTTTCAGACTCATGATAGTAAGAGCACC 1080
QY 1081 TGGCAATGATGAACGAAATATTAGGACCCATACCAACACATGATTCAGAAACACAGAA 1140
Db 1081 TGGCAATGATGAACGAAATATTAGGACCCATACCAACACATGATTCAGAAACACAGAA 1140
QY 1141 AACCAAGTATTTTCCACCATTAACAGCTAGATGAGGATGAAACACACAGTTCTGCTGTAG 1200
Db 1141 AACCAAGTATTTTCCACCATTAACAGCTAGATGAGGATGAAACACACAGTTCTGCTGTAG 1200
QY 1201 ATGTTAGAGAGCGTGCACCAACCGTTGAAGAAATTTATGCTTTGTGATGATGATGATG 1260
Db 1201 ATGTTAGAGAGCGTGCACCAACCGTTGAAGAAATTTATGCTTTGTGATGATGATGATG 1260
QY 1261 AGAACTGTTTGAAGTGGTTCGAAAGAAATGTTAGAAATATGATCCCACTCAAGAAATTA 1320
Db 1261 AGAACTGTTTGAAGTGGTTCGAAAGAAATGTTAGAAATATGATCCCACTCAAGAAATTA 1320
QY 1321 TGGATGAAGCAATGCGAGCATCTCTTCTTGTGCTTATTTAAAGAAATGAAATGGAATC 1380
Db 1321 TGGATGAAGCAATGCGAGCATCTCTTCTTGTGCTTATTTAAAGAAATGAAATGGAATC 1380
QY 1381 AGTGTCTTATATATATCTCTAGAAAGATTTCTTAAAGATGCTGTGCTCAACTTAA 1440
Db 1381 AGTGTCTTATATATATCTCTAGAAAGATTTCTTAAAGATGCTGTGCTCAACTTAA 1440
QY 1441 CATCTTAATATTTTGTAAACATTTAAATTTATTTGTACAGTTAAGTGAATATTTGTATG 1500
Db 1441 CATCTTAATATTTTGTAAACATTTAAATTTATTTGTACAGTTAAGTGAATATTTGTATG 1500
QY 1501 TTTTGTATCAATAGCATAAATTAATCTTGTAAAGCAAGTATGCTTGTGATAATGCAATGAA 1560
Db 1501 TTTTGTATCAATAGCATAAATTAATCTTGTGTAAAGCAAGTATGCTTGTGATAATGCAATGAA 1560
QY 1561 AAATTAATAATTTTCTTTTGAATTAACATTTTAAATACCTTTTGAATATCTTT 1620
Db 1561 AAATTAATAATTTTCTTTTGAATTAACATTTTAAATACCTTTTGAATATCTTT 1620
QY 1621 TGTGTCCAGTCAATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680
Db 1621 TGTGTCCAGTCAATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680
QY 1681 GATTTTTTTTTCAGTAAAGGAAATCTTGCATCTTATTTATTTTAAAGGAAATTTCTTTA 1740
Db 1681 GATTTTTTTTTCAGTAAAGGAAATCTTGCATCTTATTTATTTTAAAGGAAATTTCTTTA 1740
QY 1741 TATATCTCAAAATTTAGAACTTAACTTTTAAAGTCTTCTTCTGTAATTTGTTGAACGGGTG 1800
Db 1741 TATATCTCAAAATTTAGAACTTAACTTTTAAAGTCTTCTTCTGTAATTTGTTGAACGGGTG 1800
QY 1801 ATTATTTATTAATCTAGATAGCAGGTAAGAAACCAAACTCAGAAATGTTTACTGT 1860
Db 1801 ATTATTTATTAATCTAGATAGCAGGTAAGAAACCAAACTCAGAAATGTTTACTGT 1860

Qy	1861	TAGAA	TCTATTA	AAAA	TTTTAA	GTGTTG	TATCT	TTTTCA	TTCGG	TGATG	TCA	CGGG	TGATA	1920
Db	1861	TAGAA	TCTATTA	AAAA	TTTTAA	GTGTTG	TATCT	TTTTCA	TTCGG	TGATG	TCA	CGGG	TGATA	1920
Qy	1921	ACAGAC	ATTCA	TGGAA	AGCATG	CAGTTGT	CCATTG	CGAC	AGTTG	TGTA	TTGTTT	TAATA	AAAA	1980
Db	1921	ACCAGAC	ATTCA	TGGAA	AGCATG	CAGTTGT	CCATTG	CGAC	AGTTG	TGTA	TTGTTT	TAATA	AAAA	1980
Qy	1981	ACATAC	CACACT	TTATTT	TAAGAT	TAAATCT	TAAC	CTGGA	AAAGT	TCAG	CTTGG	AAAAAT	TGGACAT	2040
Db	1981	ACATAC	CACACT	TTATTT	TAAGAT	TAAATCT	TAAC	CTGGA	AAAGT	TCAG	CTTGG	AAAAAT	TGGACAT	2040
Qy	2041	TTCCA	AGTATG	TTTGG	TGAGT	TCACAG	ATATA	AAAAAT	TCTG	ATGAG	AGGTTT	TCAG	2100	
Db	2041	TTCCA	AGTATG	TTTGG	TGAGT	TCACAG	ATATA	AAAAAT	TCTG	ATGAG	AGGTTT	TCAG	2100	
Qy	2101	TTTTT	TAATAC	CAAGT	CTCTTA	AGGAGT	CTTA	ACA	TTGGC	CAGCAT	CTGTTT	ATCAA	ATGACA	2160
Db	2101	TTTTT	TAATAC	CAAGT	CTCTTA	AGGAGT	CTTA	ACA	TTGGC	CAGCAT	CTGTTT	ATCAA	ATGACA	2160
Qy	2161	TAAAT	ACGTAA	ACCTATA	TAAGAA	TTTAA	AGT	TTTAA	TTTAA	TAGG	CAATTT	TATG	TCTGTG	2220
Db	2161	TAAAT	ACGTAA	ACCTATA	TAAGAA	TTTAA	AGT	TTTAA	TTTAA	TAGG	CAATTT	TATG	TCTGTG	2220
Qy	2221	TCTT	ACGGG	AGAA	GAGG	ATTTG	ATTG	GAAG	CAGAT	TTTGG	GAA	GAAGT	TCTG	2280
Db	2221	TCTT	ACGGG	AGAA	GAGG	ATTTG	ATTG	GAAG	CAGAT	TTTGG	GAA	GAAGT	TCTG	2280
Qy	2281	TTCC	AGAA	TTTAA	TGAT	TGGT	TACATA	AAATCT	TTTTC	GACT	TCAG	AAAA	AAAA	2340
Db	2281	TTCC	AGAA	TTTAA	TGAT	TGGT	TACATA	AAATCT	TTTTC	GACT	TCAG	AAAA	AAAA	2340
Qy	2341	AACAAA	AAAA	AAAA	AAAC	2354								
Db	2341	AACAAA	AAAA	AAAA	AAAC	2354								
RESULT 3														
AR534243														
LOCUS	AR534243													
DEFINITION	Sequence 1 from patent US 6733978.													
ACCESSION	AR534243													
VERSION	AR534243.1													
KEYWORDS	GI:53924333													
SOURCE	Unknown.													
ORGANISM	Unclassified.													
REFERENCE	1 (bases 1 to 2354)													
AUTHORS	Yan, C., Ye, J., Ketchum, K.A., Di Francesco, V. and Beasley, E.M.													
TITLE	Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins, and uses thereof													
JOURNAL	Patent: US 6733978-A 11-MAY-2004;													
FEATURES	Location/Qualifiers													
source	1..2354													
ORIGIN	/organism="unknown"													
	/mol_type="genomic DNA"													
Query Match	100.0%;	Score	2354;	DB	6;	Length	2354;							
Best Local Similarity	100.0%;	Pred.No.	0;											
Matches 2354;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;						
Qy	1	GCCAGCTGGGGTTACTTTTAA	AAAAACATGCTCCATGTC	GCATCCCTCTTTGAAGCTTCGC	ACT	60								
Db	1	GCCAGCTGGGGTTACTTTTAA	AAAAACATGCTCCATGTC	GCATCCCTCTTTGAAGCTTCGC	ACT	60								
Qy	61	CTGTTTGAAGAGGACACATC	ATCCAGTCATTTATTTAG	AAAGCAAGCTCCTTCAATGAGCG	AG	120								
Db	61	CTGTTTGAAGAGGACACATC	ATCCAGTCATTTATTTAG	AAAGCAAGCTCCTTCAATGAGCG	AG	120								
Qy	121	ATTATCGGGA	CCGGAGATACGTTG	CAAT										


```
Qy 601 TAGAATGGTTTGATCATCATGGTCAATGTTCTGTATGTTGTTGAACTACTGGGACTTAGTA 660
Db 601 TAGAATGGTTTGATCATCATGGTCAATGTTCTGTATGTTGTTGAACTACTGGGACTTAGTA 660
Qy 661 CTTACGATTTTCATTAAAGAAAACAGCTTTCTGCCATTTCAAATTTGACCACATCAGGCAGA 720
Db 661 CTTACGATTTTCATTAAAGAAAACAGCTTTCTGCCATTTCAAATTTGACCACATCAGGCAGA 720
Qy 721 TGGCGTATCAGATCTGCCAGTCAATAAATTTTACATCATATAAATAAATTAACCCATACAG 780
Db 721 TGGCGTATCAGATCTGCCAGTCAATAAATTTTACATCATATAAATAAATTAACCCATACAG 780
Qy 781 ATCTGAAGCCCTGAATAATATTTTGTGTGAAGTCTGACTAGTCAAAATATAATTTCTA 840
Db 781 ATCTGAAGCCCTGAATAATATTTTGTGTGAAGTCTGACTAGTCAAAATATAATTTCTA 840
Qy 841 AAATGAACGATGATGAACGACACACTGAAAACACAGATATCAAAAGTTGTTCACCTTTGGAA 900
Db 841 AAATGAACGATGATGAACGACACACTGAAAACACAGATATCAAAAGTTGTTCACCTTTGGAA 900
Qy 901 GTGCAACGTATGATGATGAACATCACTACTTTTGGTGTCTACCCGGCAGCTACAGAGCTC 960
Db 901 GTGCAACGTATGATGATGAACATCACTACTTTTGGTGTCTACCCGGCAGCTACAGAGCTC 960
Qy 961 CCGAGGTCAATTTGGCTTTAGGTTGGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCA 1020
Db 961 CCGAGGTCAATTTGGCTTTAGGTTGGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCA 1020
Qy 1021 TTCTTATTTGAATATTTACCTTGGTTTACAGTCTTTTACAGTCTCATGATAGTAAGAGCACC 1080
Db 1021 TTCTTATTTGAATATTTACCTTGGTTTACAGTCTTTTACAGTCTCATGATAGTAAGAGCACC 1080
Qy 1081 TGGCAATCATGAAAGCAATATTTAGACCCATACCAACACATCATGATTCAGAAAAACAAGAA 1140
Db 1081 TGGCAATCATGAAAGCAATATTTAGACCCATACCAACACATCATGATTCAGAAAAACAAGAA 1140
Qy 1141 AACGCAAGTATTTTCAACATTAACAGCTAGATTTGGATGAAACACAGATTTCTGCTGGTAGAT 1200
Db 1141 AACGCAAGTATTTTCAACATTAACAGCTAGATTTGGATGAAACACAGATTTCTGCTGGTAGAT 1200
Qy 1201 ATGTTAGAGAGCTGCAACCGTTGAGGAAATTTATGCTTTTTCATGATCAAGAAACATG 1260
Db 1201 ATGTTAGAGAGCTGCAACCGTTGAGGAAATTTATGCTTTTTCATGATCAAGAAACATG 1260
Qy 1261 AGAAACTGTTTGACCTGTTTCCGAGAAATTTAGAAATATGATCCAACTCAAGAAATTAACCT 1320
Db 1261 AGAAACTGTTTGACCTGTTTCCGAGAAATTTAGAAATATGATCCAACTCAAGAAATTAACCT 1320
Qy 1321 TGGATGAAGCATTCAGCATCTTTCTTTGACTTATTAATAAAGAAATGAAATGGGAATC 1380
Db 1321 TGGATGAAGCATTCAGCATCTTTCTTTGACTTATTAATAAAGAAATGAAATGGGAATC 1380
Qy 1381 AGTGGTCTTACTATATATCTTCTAGAGAGATTTACTTTAAGACGTGTGTCACTCAACTAAA 1440
Db 1381 AGTGGTCTTACTATATATCTTCTAGAGAGATTTACTTTAAGACGTGTGTCACTCAACTAAA 1440
Qy 1441 CATTCATAATATTTTGTAAACATTAATTTTGTACAGTTAAGTGTAAATATTTGTATG 1500
Db 1441 CATTCATAATATTTTGTAAACATTAATTTTGTACAGTTAAGTGTAAATATTTGTATG 1500
Qy 1501 TTTTGTATCAATAGCATAAATTAACCTTTGTAAGCAAGTATGCTTTGTGATTAATGCAATGAA 1560
Db 1501 TTTTGTATCAATAGCATAAATTAACCTTTGTAAGCAAGTATGCTTTGTGATTAATGCAATGAA 1560
Qy 1561 AAATTAATAATTAATTTTCTTTTGAATATACCAATTTTAAATACCTTTTGAATATCTCTT 1620
Db 1561 AAATTAATAATTAATTTTCTTTTGAATATACCAATTTTAAATACCTTTTGAATATCTCTT 1620
Qy 1621 TGTGTCCAGTGATTAATGATGATCTTGGCTTTTGTACATGAGGAGTCACTCTCTGAAGT 1680
Db 1621 TGTGTCCAGTGATTAATGATGATCTTGGCTTTTGTACATGAGGAGTCACTCTCTGAAGT 1680
Qy 1681 GATTTTTTTTGAGTAAAGGAAATCTTGACTACTTTTATTTTCTTAAGGAATATTTCTTTTA 1740
```

```
Db 1681 GATTTTTTTTGAGTAAAGGAAATCTTGACTACTTTATATTTCTTAAGGAATATTTCTTTTA 1740
Qy 1741 TATACTTCAAATTTAGAACTTAACTTTAAAGTTTTCCTCTGTAATTTGTTGAACGGGTG 1800
Db 1741 TATACTTCAAATTTAGAACTTAACTTTAAAGTTTTCCTCTGTAATTTGTTGAACGGGTG 1800
Qy 1801 ATTATTATTAACCTCTAGATAAGCAGGTACTAGAAAACCAAACTCAGAAAATGTTTACTGT 1860
Db 1801 ATTATTATTAACCTCTAGATAAGCAGGTACTAGAAAACCAAACTCAGAAAATGTTTACTGT 1860
Qy 1861 TAGAATTTCTATTAATTTTAAAGTTTGTATTTCTTTTTCATTTGGTGTGATGTCAGGGTGATA 1920
Db 1861 TAGAATTTCTATTAATTTTAAAGTTTGTATTTCTTTTTCATTTGGTGTGATGTCAGGGTGATA 1920
Qy 1921 ACCAGACATTCATGGAAGGCAATGTCAGTTTGTCCATTTGTGCACAGTTTGTTTAATAAACC 1980
Db 1921 ACCAGACATTCATGGAAGGCAATGTCAGTTTGTCCATTTGTGCACAGTTTGTTTAATAAACC 1980
Qy 1981 ACATACACACTTTTATTTAAGATTAATAATCTAACTGGAAGTCAAGCTTGGAAAATGGACAT 2040
Db 1981 ACATACACACTTTTATTTAAGATTAATAATCTAACTGGAAGTCAAGCTTGGAAAATGGACAT 2040
Qy 2041 TTCCAAGTATGTTTGGTGTGAGTCTTAACATTTGGCCAGCATCTGTTTATCAAAATGACA 2100
Db 2041 TTCCAAGTATGTTTGGTGTGAGTCTTAACATTTGGCCAGCATCTGTTTATCAAAATGACA 2100
Qy 2101 TTTTAAATACCAAGTCTTAGGAGTCTTAACATTTGGCCAGCATCTGTTTATCAAAATGACA 2160
Db 2101 TTTTAAATACCAAGTCTTAGGAGTCTTAACATTTGGCCAGCATCTGTTTATCAAAATGACA 2160
Qy 2161 TAAATACGTAAACCTATAAAGATTTAAGTTTAAATTTAGGCAATTTATGTCGTGATAAT 2220
Db 2161 TAAATACGTAAACCTATAAAGATTTAAGTTTAAATTTAGGCAATTTATGTCGTGATAAT 2220
Qy 2221 TCTTACGGGAGAAAGAGGATTTGTTGGAAGCAGATTTGGGAAGAAAGTCTCTGTAAT 2280
Db 2221 TCTTACGGGAGAAAGAGGATTTGTTGGAAGCAGATTTGGGAAGAAAGTCTCTGTAAT 2280
Qy 2281 TTCCAGAAATTTAATTTGATTTGATTTACATAAATTTTGGACTTTCAGAAAAAATAAAAA 2340
Db 2281 TTCCAGAAATTTAATTTGATTTGATTTACATAAATTTTGGACTTTCAGAAAAAATAAAAA 2340
Qy 2341 AACAAAAAATAAAC 2354
Db 2341 AACAAAAAATAAAC 2354
```

```
RESULT 5
AR077239
LOCUS AR077239 2446 bp DNA linear PAT 31-AUG-2000
DEFINITION Sequence 9 from patent US 5962232.
ACCESSION AR077239
VERSION AR077239.1 GI:10003985
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2446)
AUTHORS Bandman,O., Hillman,J.L., Lal,P., Akerblom,I.E., Shah,P.,
Corley,N.C. and Guegler,K.J.
TITLE Protein Kinase molecules
JOURNAL Patent: US 5962232-A 9 05-OCT-1999;
FEATURES Location/Qualifiers
source 1. 2446
/organism="unknown"
/mol_type="unassigned DNA"
```

```
ORIGIN
Query Match 87.2%; Score 2052.6; DB 6; Length 2446;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 2155; Conservative 0; Mismatches 4; Indels 90; Gaps 1;
```


QY	85	GTCAATATTAGAACCAAGTCCTTGAATGACGAGATATTACGGACCGGAGATACGTTG	144	Db	1278	AGCTAGATTGGGATGAAACACAGTTCTGCTGTAGATATGTTAGGAGACGCTGCAAAACCGT	1337
Db	288	GTCAATATTAGAACCAAGTCCTTGAATGACGAGATATTACGGACCGGAGATACGTTG	347	QY	1225	TGAAGGAATTTATGCTTGTGTCATGATGAAGAACATGAGAAAATGTTTGCACCTGGTTCGAA	1284
QY	145	ACGAATACAGGAATGACCTACTGTTGAAGGATATGTTCTAGACATTTATCACAGAGACATTG	204	Db	1338	TGAAGGAATTTATGCTTGTGTCATGATGAAGAACATGAGAAAATGTTTGCACCTGGTTCGAA	1397
Db	348	ACGAATACAGGAATGACCTACTGTTGAAGGATATGTTCTAGACATTTATCACAGAGACATTG	407	QY	1385	GAATGTTAGGAATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCAATTCGACGATCCTT	1444
QY	205	AAAGCGGATATCGAATCCATCGCAGTAAATCTTCAGTCCGACGAGGAGAGACGATCCCTA	264	Db	1398	GAATGTTAGGAATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCAATTCGACGATCCTT	1457
Db	408	AAAGCGGATATCGAATCCATCGCAGTAAATCTTCAGTCCGACGAGGAGAGACGATCCCTA	467	QY	1345	TCCTTGACTTTATTAAGAAAGAAATGAAATGGAATCAGTGGCTCTTACTATATATCTTCTCT	1404
QY	265	AAAGGAACGCAATAGACATCTGTTTCAAGTCATGATGATCGTTTCGAGAGCCACCGAAGGA	324	Db	1458	TCCTTGACTTTATTAAGAAAGAAATGAAATGGAATCAGTGGCTCTTACTATATATCTTCTCT	1517
Db	468	AAAGGAACGCAATAGACATCTGTTTCAAGTCATGATGATCGTTTCGAGATG	515	QY	1405	AGAAAGAGATTACTTAAGACTGTGTCAGTCAACTAAACATCTTAAATATTTTGTGAACATTT	1464
QY	325	AAAGATCCAGGATATAGAGGATGATGAGGAGGTCACTGATCTGTCAAAAGTGGAGACG	384	Db	1518	AGAAGAGATTACTTAAGACTGTGTCAGTCAACTAAACATCTTAAATATTTTGTGAACATTT	1577
Db	516	-----	515	QY	1465	AAATTAATTTGTGACAGTTAAGTGAATATTTGTATGTTTGTATCAATPAGCATATTAAC	1524
QY	385	TTCTAAGACGAAGATPATGAATCGTGGACACCTTTGGGTGAAGGAGCCTTTGGCAAAAGTTG	444	Db	1578	AAATTAATTTGTGACAGTTAAGTGAATATTTGTATGTTTGTATCAATPAGCATATTAAC	1637
Db	516	-----AATCCGTGGACACTTTTGGGTGAAGGAGCCTTTGGCAAAAGTTG	557	QY	1525	TTGTTAAGCAAGTATGCTTGTGATAATGCAATAGAAAATTAATAATTTTCTTTCTTTT	1584
QY	445	TAGAGTGCATTTGATCGGCATGAGATGGCATGCGATGTAAGCAATGTAATAAATG	504	Db	1638	TTGTTAAGCAAGTATGCTTGTGATAATGCAATAGAAAATTAATAATTTTCTTTCTTTT	1697
Db	558	TAGAGTGCATTTGATCGGCATGAGATGGCATGCGATGTAAGCAATGTAATAAATG	617	QY	1585	GAAATTAACCATTTTAAATACCTTTGAAATATCTTGTGTCAGTGAATAATGTAATG	1644
QY	505	TAGGCGGTTACCGTGAAGCAGCTCGTTCCAGAAATCCAAAGTATTAGAGCATTAAATAGTA	564	Db	1698	GAAATTAACCATTTTAAATACCTTTGAAATATCTTGTGTCAGTGAATAATGTAATG	1757
Db	618	TAGGCGGTTACCGTGAAGCAGCTCGTTCCAGAAATCCAAAGTATTAGAGCATTAAATAGTA	677	QY	1645	ATCTGCGCTTTTGTACATGAGGTCACCTCTGAAAGTGAATTTTGTAGTAAAAAGGAAAT	1704
QY	565	CTGATCCCAATAGTGTCTTCCGATGTCGAGATGCTAGAAATGTTTGCATCATCATGTC	624	Db	1758	ATCTGCGCTTTTGTACATGAGGTCACCTCTGAAAGTGAATTTTGTAGTAAAAAGGAAAT	1817
Db	678	CTGATCCCAATAGTGTCTTCCGATGTCGAGATGCTAGAAATGTTTGCATCATCATGTC	737	QY	1705	CTTGACTACTTTATATCTTAAAGGAATATTTCTTTATATATCTTCAAAATTTAGAACTTAAC	1764
QY	625	ATGTTTGTATTTGTTGAACTAATGCGGATCTAGTCTTACATTTCAATTAAGAAACA	684	Db	1818	CTTGACTACTTTATATCTTAAAGGAATATTTCTTTATATATCTTCAAAATTTAGAACTTAAC	1877
Db	738	ATGTTTGTATTTGTTGAACTAATGCGGATCTAGTCTTACATTTCAATTAAGAAACA	797	QY	1765	TTTAAAGTCTTCTCTGTAATTTGTAAGCGGTGATTAATTAACCTCTAGATAAGCA	1824
QY	685	GCTTCTGCCATTTCAAAATGACCAATGACCAATGACCAATGACCAATGACCAATGACCAAT	744	Db	1878	TTTAAAGTCTTCTCTGTAATTTGTAAGCGGTGATTAATTAACCTCTAGATAAGCA	1937
Db	798	GCTTCTGCCATTTCAAAATGACCAATGACCAATGACCAATGACCAATGACCAATGACCAAT	857	QY	1825	GGTACTAGAAACCAAACTCAGAAAATGTTTACTGTTAGAAATCTTATTAATTTTAAAGTG	1884
QY	745	TAAATTTTATCATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	804	Db	1938	GGTACTAGAAACCAAACTCAGAAAATGTTTACTGTTAGAAATCTTATTAATTTTAAAGTG	1997
Db	858	TAAATTTTATCATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	917	QY	1885	TTGTAATCTTTTCTTCAATTTGGGTGATGTCAGGTGATTAACAGACATTTATGGAAGGCAATG	1944
QY	805	TTGTGAAGTCTGACTATGATGATCAATATTAATTTCAAAATGAAACGTCATGAAACGACAC	864	Db	1998	TTGTAATCTTTTCTTCAATTTGGGTGATGTCAGGTGATTAACAGACATTTATGGAAGGCAATG	2057
Db	918	TTGTGAAGTCTGACTATGATGATCAATATTAATTTCAAAATGAAACGTCATGAAACGACAC	977	QY	1945	CAGTTTGTCCATTTGTGACAGTTTGTAAATAAAACCATACACATTTTATTTAAAGATTA	2004
QY	865	TGAAAACACAGATATCAAGTTGTTGACTTTGAGTCCCGAGTCCCGAGTCCCGAGTCCCGAGT	924	Db	2058	CAGTTTGTCCATTTGTGACAGTTTGTAAATAAAACCATACACATTTTATTTAAAGATTA	2117
Db	978	TGAAAACACAGATATCAAGTTGTTGACTTTGAGTCCCGAGTCCCGAGTCCCGAGTCCCGAGT	1037	QY	2005	AAATCTAACTGAAAGTCAGCTTGGAAAATGAGCAATTTTCCAAAGTATGTTTGGTGAAGTCAC	2064
QY	925	ACAGTACTTTGGTGTCTACCGGCACTACAGAGTCCCGAGTCCCGAGTCCCGAGTCCCGAGT	984	Db	2118	AAATCTAACTGAAAGTCAGCTTGGAAAATGAGCAATTTTCCAAAGTATGTTTGGTGAAGTCAC	2177
Db	1038	ACAGTACTTTGGTGTCTACCGGCACTACAGAGTCCCGAGTCCCGAGTCCCGAGTCCCGAGT	1097	QY	2065	AGATATAAAATAGAAATTCGATGAGAGGTTTTCAGTTTAAATACCAAGTCTTTAGAG	2124
QY	985	GGTCTCAGCCTTTGATGTTTGGAGCAATGAGTGTGATTTTATTTGAATATTTACCTTGGTT	1044	Db	2178	AGATATAAAATAGAAATTCGATGAGAGGTTTTCAGTTTAAATACCAAGTCTTTAGAG	2237
Db	1098	GGTCTCAGCCTTTGATGTTTGGAGCAATGAGTGTGATTTTATTTGAATATTTACCTTGGTT	1157	QY	2125	TCCTTAAACATTTGGCCGAGCATCTGTTTATCAAAATGACATAAAATACGTAACCTTAAAGAAAT	2184
QY	1045	TCAGAGTCTTTGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1104	Db	2238	TCCTTAAACATTTGGCCGAGCATCTGTTTATCAAAATGACATAAAATACGTAACCTTAAAGAAAT	2297
Db	1158	TCAGAGTCTTTGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1217	QY	2185	AAGTTTATTAATAGGCAATTTATGCTGCTGATTAATCTTACGGGAGAGAGGATTTGA	2244
QY	1105	GACCCATACCAACACATGATTTAGAAAACGCAAGTATTTTACCATTAAC	1164	Db	2298	AAGTTTATTAATAGGCAATTTATGCTGCTGATTAATCTTACGGGAGAGAGGATTTGA	2357
Db	1218	GACCCATACCAACACATGATTTAGAAAACGCAAGTATTTTACCATTAAC	1277	QY	2245	TTGAAAGCAGTTTGGGAGAGAAAGTGTGCTGAAATTTTCCAGAAATTTTAAATTTGAATGGTTA	2304
QY	1165	AGCTAGATTGGGATGAACACAGTTCTGCTGTGATATGTTAGGAGACGCTGCAAAACCGT	1224				

Db 2358 TTGGAAGCAGTTTGGGAAGAAAGTGCTGCTGAAATTTCCAGAAATTAATTAATGGTTA 2417

Qy 2305 CATAACTTTTGAAGCTTTCAGAAAAA 2333

Db 2418 CATAACTTTTGAAGCTTTCAGAAAAA 2446

RESULT 6

BD107903

LOCUS BD107903 2446 bp DNA linear PAT 18-SEP-2002

DEFINITION Protein kinase molecules.

ACCESSION BD107903

VERSION BD107903.1 GI:23202721

KEYWORDS JP 2002501749-A/3

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Bandman,O., Hillman,J.L., Lal,P., Akerblom,I.E., Shah,P., Corley,N.C. and Guegler,K.J.

TITLE Protein kinase molecules

JOURNAL INCYTE PHARMACEUTICALS INC

COMMENT Patent: JP 2002501749-A 3 22-JAN-2002; OS Homo sapiens (human) PN JP 2002501749-A/3 PD 22-JAN-2002 PF 12-JAN-1999 JP 2000529440 PR 30-JAN-1998 US 09/016000 PI OLGA BANDMAN, JENNIFER L HILLMAN, PREETI LAL, INGRID E AKERBLOM, PI PURVI SHAH, PI NEIL C CORLEY, KARL J GUEGLER PC C12N15/09,A61K38/45,A61K45/00,C07K16/40,C12N1/15,C12N1/19, PC C12N1/21, PC C12N5/10,C12N9/12,C12Q1/68,G01N33/68,C12N15/00,A61K37/52,C12N5/PC 00 CC 339963 Location/Qualifiers FH Key I. .2446 /organism='Homo sapiens (human)'. FT source Location/Qualifiers 1. .2446 /organism='Homo sapiens' /mol_type='genomic DNA' /db_xref='taxon:9606'

FEATURES

source

ORIGIN

Query Match 87.2%; Score 2052.6; DB 6; Length 2446;

Best Local Similarity 95.8%; Pred. No. 0;

Matches 2155; Conservative 0; Mismatches 4; Indels 90; Gaps 1;

Qy 85 GTCATTATTGAAGCAAGGCTTTGAATGAGCGAGATTATCGGACCGGAGATACGTTG 144

Db 288 GTCATTATTGAAGCAAGGCTTTGAATGAGCGAGATTATCGGACCGGAGATACGTTG 347

Qy 145 ACGATACAGGANTGACTGTGAGGATATGTTCCCTAGACNTTATCACAGACATTG 204

Db 348 ACGATACAGGANTGACTGTGAGGATATGTTCCCTAGACNTTATCACAGACATTG 407

Qy 205 AAAGCGGTATCGAATCCACTGCGATAATCTTCAGTCCGAGGAGAGAGAGAGAGAGAG 264

Db 408 AAAGCGGTATCGAATCCACTGCGATAATCTTCAGTCCGAGGAGAGAGAGAGAGAG 467

Qy 265 AAAGGAGCGCAATAGACACTGTTCAAGTCATCATGTCAGTTCAGTTCGAGAGAGAGAG 324

Db 468 AAAGGAGCGCAATAGACACTGTTCAAGTCATCATGTCAGTTCAGTTCGATGA----- 515

Qy 325 AAAGATCCAGGAGTATAGAGGATGATGAGGGGTCACTGATCTGTCAAAGTGGAGAG 384

Db 516 ----- 515

Qy 385 TTCTAAGAGCAAGATATGAAATCGTGGACACTTTTGGGTGAAGAGGCCCTTTGGCAAAAGTTG 444

Db 516 -----AATCCGTGGACACTTTTGGGTGAAGGAGCCCTTTGGCAAAAGTTG 557

Qy 445 TAGAGTGCATTGATCATGGCATGATGGCATGATGGCATGATGGCATGATGGCATGATGGCATG 504

Db 558 TAGAGTGCATTGATCATGGCATGATGGCATGATGGCATGATGGCATGATGGCATGATGGCATG 617

Qy 505 TAGGCCGTTACCGTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTA 564

Db 618 TAGGCCGTTACCGTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTA 577

Qy 565 CTGATCCCAATAGTGTCTTCOGATGTGTCCAGATGCTAGAAATGGTTGATCATCATGTC 624

Db 678 CTGATCCCAATAGTGTCTTCOGATGTGTCCAGATGCTAGAAATGGTTGATCATCATGTC 737

Qy 625 ATGTTGTATGTTGTTGAACACTACTGGGACTTAGTACTTACGATTTCATTAAAGAAACA 684

Db 738 ATGTTGTATGTTGTTGAACACTACTGGGACTTAGTACTTACGATTTCATTAAAGAAACA 797

Qy 685 GCTTTCGCCAATTTCAAATTCACCATCATCAGGAGATGGCGTATCAGATCTGCCAGTCAA 744

Db 798 GCTTTCGCCAATTTCAAATTCACCATCATCAGGAGATGGCGTATCAGATCTGCCAGTCAA 857

Qy 745 TAAATTTTACATCATAATAATAAACCATAAGATCTGAAAGCTGAAATATTTTGT 804

Db 858 TAAATTTTACATCATAATAATAAACCATAAGATCTGAAAGCTGAAATATTTTGT 917

Qy 805 TTGTGAAGTCTGACTATGTAGTCAATATATTTCTAAATGAAACGCTGATGATGATGATGATG 864

Db 918 TTGTGAAGTCTGACTATGTAGTCAATATATTTCTAAATGAAACGCTGATGATGATGATGATG 977

Qy 865 TGAATAACACAGATATCAAAAGTGTGCTGCTTGAAGTGAACGCTGATGATGATGATGATGATG 924

Db 978 TGAATAACACAGATATCAAAAGTGTGCTGCTTGAAGTGAACGCTGATGATGATGATGATGATG 1037

Qy 925 ACAGTACTTTGGTGTCTACCGGCACATACAGAGTCCCGAGGTCAATTTTGGCTTTAGGTT 984

Db 1038 ACAGTACTTTGGTGTCTACCGGCACATACAGAGTCCCGAGGTCAATTTTGGCTTTAGGTT 1097

Qy 985 GGTCTCAGCTTTGTGATGTTTGGAGCATAGTTCGATTTCTTATGAAATATACCTTTGGTT 1044

Db 1098 GGTCTCAGCTTTGTGATGTTTGGAGCATAGTTCGATTTCTTATGAAATATACCTTTGGTT 1157

Qy 1045 TCACAGTCTTTTCAGACTCATGATAGTAAAGACACCTGGCAATGATGAAAGCAATATTAG 1104

Db 1158 TCACAGTCTTTTCAGACTCATGATAGTAAAGACACCTGGCAATGATGAAAGCAATATTAG 1217

Qy 1105 GACCCATACCAACACATGATTCAGAAAAACAAGAAACGCAAGTATTTTCCACCAATACC 1164

Db 1218 GACCCATACCAACACATGATTCAGAAAAACAAGAAACGCAAGTATTTTCCACCAATACC 1277

Qy 1165 AGCTAGATTGGGATGAACACAGTTCCTGCTGTAGATATGTTAGGAGACGCTGCAACCGT 1224

Db 1278 AGCTAGATTGGGATGAACACAGTTCCTGCTGTAGATATGTTAGGAGACGCTGCAACCGT 1337

Qy 1225 TGAAGGAATTTATGCTTTGTCATGATGAAGAAACATGAGAAACCTGTTGACCTGGTTGCA 1284

Db 1338 TGAAGGAATTTATGCTTTGTCATGATGAAGAAACATGAGAAACCTGTTGACCTGGTTGCA 1397

Qy 1285 GAATGTTAGAAATATGATCCAACTCAAAGAAATTAACCTTGATGAAGCATGTCAGCATCCT 1344

Db 1398 GAATGTTAGAAATATGATCCAACTCAAAGAAATTAACCTTGATGAAGCATGTCAGCATCCT 1457

Qy 1345 TCTTTGACTTATTAAGAAAGAAATGAATGGAAATCAGTGGCTTACTATATCTCTCT 1404

Db 1458 TCTTTGACTTATTAAGAAAGAAATGAATGGAAATCAGTGGCTTACTATATCTCTCTCT 1517

Qy 1405 AGAAGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATTTTGTAAACATT 1464

Db 1518 AGAAGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATTTTGTAAACATT 1577

Qy 1465 AAATTTTGTACAGTTAAGTGAATATTTGTATGTTTGTATCAATAGCATTAATTAAC 1524

Db 1578 AAATTAATTTGTACAGTTAAGTGTAAATATTGTATGTTTGTATCAATAGCATATAATTAAC 1637
QY 1525 TTGTTAAGCAAGTATGGTCTTGATAATGCATTAGAAAAATTAATAATTTCTTTT 1584
Db 1638 TTGTTAAGCAAGTATGGTCTTGATAATGCATTAGAAAAATTAATAATTTCTTTT 1697
QY 1595 GAAATACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATAATGATGATG 1644
Db 1698 GAAATACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATAATGATGATG 1757
QY 1645 ATCTTGCCCTTTGTACATGAGAGTCACTCTCAAGTGATTTTGTGATGAAGAAAT 1704
Db 1758 ATCTTGCCCTTTGTACATGAGAGTCACTCTCAAGTGATTTTGTGATGAAGAAAT 1817
QY 1705 CTGTGACTACTTTATATTTCTTAAAGGAATATCTTTATATATCTTCAAAATTAGAACTTAAC 1764
Db 1818 CTGTGACTACTTTATATTTCTTAAAGGAATATCTTTATATATCTTCAAAATTAGAACTTAAC 1877
QY 1765 TTTAAAGTTTCTTCTGTAAATTTGTGAACGGGTGATTAATTAATCTAGATAAGCA 1824
Db 1878 TTTAAAGTTTCTTCTGTAAATTTGTGAACGGGTGATTAATTAATCTAGATAAGCA 1937
QY 1825 GGTACTAGAAACCAAACTCAGAAATGTTTACTGTAGAAATCTTATTAATTTTAAAGT 1884
Db 1938 GGTACTAGAAACCAAACTCAGAAATGTTTACTGTAGAAATCTTATTAATTTTAAAGT 1997
QY 1885 TTGTATTTCTTTTCAATGGGTGATGTCAGGGGTGATAACAGACATTCATGGAAGGCATG 1944
Db 1998 TTGTATTTCTTTTCAATGGGTGATGTCAGGGGTGATAACAGACATTCATGGAAGGCATG 2057
QY 1945 CAGTTTGTCCATTTGTGACATTTGTTTAAATAAAACCATACACATCACTTTATTAAGATTA 2004
Db 2058 CAGTTTGTCCATTTGTGACATTTGTTTAAATAAAACCATACACATCACTTTATTAAGATTA 2117
QY 2005 AAATCTAACTGGAAGTCACTGTGAAATGCAATCTTCCAGATGTTTGGTGAATCAC 2064
Db 2118 AAATCTAACTGGAAGTCACTGTGAAATGCAATCTTCCAGATGTTTGGTGAATCAC 2177
QY 2065 AGATAATAAATAGAAATCTGATGAGAGTTTCAGTTTAAATCAAGTCTTTAGGAG 2124
Db 2178 AGATAATAAATAGAAATCTGATGAGAGTTTCAGTTTAAATCAAGTCTTTAGGAG 2237
QY 2125 TCTTAACATGCGCAGCATCTGTTTATCAATGACATAAATACGTAACCTATAGAAT 2184
Db 2238 TCTTAACATGCGCAGCATCTGTTTATCAATGACATAAATACGTAACCTATAGAAT 2297
QY 2185 AAGTTTATTAATAGCAATTTATGCTGTGATTAATCTTACGGGAGAAAGGATTTGA 2244
Db 2298 AAGTTTATTAATAGCAATTTATGCTGTGATTAATCTTACGGGAGAAAGGATTTGA 2357
QY 2245 TTGGAAGCAGTTTGGGAAGAAAGTCTGCTGAAATTTCCAGAAATTTAATGATGTTGA 2304
Db 2358 TTGGAAGCAGTTTGGGAAGAAAGTCTGCTGAAATTTCCAGAAATTTAATGATGTTGA 2417
QY 2305 CATAACTTTTGTACTTCAGAAAAA 2333
Db 2418 CATAACTTTTGTACTTCAGAAAAA 2446

RESULT 7
AX056410
LOCUS AX056410 2488 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 54 from Patent WO0073469.
ACCESSION AX056410
VERSION AX056410.1 GI:12229117
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Florman,G.D.; Martinez,R.; Whyte,D. and Sudersanam,S.
REFERENCE
AUTHORS
TITLE Protein kinases

JOURNAL Patent: WO 0073469-A 54 07-DEC-2000;
Sugen, Inc. (US)
FEATURES Location/Qualifiers
source 1..2488
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 87.0%; Score 2048; DB 6; Length 2488;
Best Local Similarity 95.9%; Pred No. 0;
Matches 2148; Conservative 0; Mismatches 1; Indels 90; Gaps 1;
QY 85 GTCATTATTTAGAACCAAGGTCCTTGAATGAGCGAGATTTATCGGACCGGAGATACGTTG 144
Db 338 GTCATTATTTAGAACCAAGGTCCTTGAATGAGCGAGATTTATCGGACCGGAGATACGTTG 397
QY 145 ACGAATACAGGAATGACTACTGTGAGGATATGTTCTTAGACATTTATCACAGACATTTG 204
Db 398 ACGAATACAGGAATGACTACTGTGAGGATATGTTCTTAGACATTTATCACAGACATTTG 457
QY 205 AAAGCGGTATCGAATCCACTGCAGTAAATCTTCACTCCGACGAGGAGAACGACGTCCTA 264
Db 458 AAAGCGGTATCGAATCCACTGCAGTAAATCTTCACTCCGACGAGGAGAACGACGTCCTA 517
QY 265 AAAGGAACGCAATAGACACATGTTTCAAGTCATCATGTCAGTTCGAAAGAGCCACCGAAGCA 324
Db 518 AAAGGAACGCAATAGACACATGTTTCAAGTCATCATGTCAGTTCGAAAGAGCCACCGAAGCA 561
QY 325 AAAGATCCAGGAGTATAGAGATGATGAGAGGGTCACTGATCTGTCAAAATGTGAGAGC 384
Db 562 ----- 561
QY 385 TTCTAAGAGCAAGATATGAATCGTGGACACTTTGGGTGAAGGAGCCCTTTCGCAAGTTG 444
Db 562 -----NATGAATCGTGGACACTTTGGGTGAAGGAGCCCTTTCGCAAGTTG 607
QY 445 TAGAGTGCATTTGATCATGCGCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTC 504
Db 608 TAGAGTGCATTTGATCATGCGCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTC 667
QY 505 TAGGCGTTTACCGTGAAGCAGCTCGTTCAAGAAATCCAAATATAGAGCACTTAAATATGTA 564
Db 668 TAGGCGTTTACCGTGAAGCAGCTCGTTCAAGAAATCCAAATATAGAGCACTTAAATATGTA 727
QY 565 CTGATCCCAATAGTCTCTTCGATGTCGTCAGAGTCTAGAAATGTTTTCATCATCATGTC 624
Db 728 CTGATCCCAATAGTCTCTTCGATGTCGTCAGAGTCTAGAAATGTTTTCATCATCATGTC 787
QY 625 ATGTTTGTATTTGTGTTTGAACCTACTGGGACTTTAGTACTTTAGTACTTTCAATTAAGAAACA 684
Db 788 ATGTTTGTATTTGTGTTTGAACCTACTGGGACTTTAGTACTTTCAATTAAGAAACA 847
QY 685 GCTTTCTGCCATTTCAAATTCACCAATCAGGAGATGCGGTATCAGATCTGCCAGTCAA 744
Db 848 GCTTTCTGCCATTTCAAATTCACCAATCAGGAGATGCGGTATCAGATCTGCCAGTCAA 907
QY 745 TAAATTTTACATCATATAAATTAACCCATACAGATCTGAAGCTGAAATATTTTGT 804
Db 908 TAAATTTTACATCATATAAATTAACCCATACAGATCTGAAGCTGAAATATTTTGT 967
QY 805 TTGTGAAGTCTGACTATGTCATCAAAATATAATTTCTTAAATGAAACGTCATGAACCCACAC 864
Db 968 TTGTGAAGTCTGACTATGTCATCAAAATATAATTTCTTAAATGAAACGTCATGAACCCACAC 1027
QY 865 TGAATAACACAGATATCAAAAGTGTGTTGATCTTTGGAAGTGCACCGTATGATGATGAACATC 924
Db 1028 TGAATAACACAGATATCAAAAGTGTGTTGATCTTTGGAAGTGCACCGTATGATGATGAACATC 1087
QY 925 ACAGTACTTTGGTGTCTACCGGCACTACAGAGTCCCGAGGTCAATTTTGGCTTTAGGTT 984
Db 1088 ACAGTACTTTGGTGTCTACCGGCACTACAGAGTCCCGAGGTCAATTTTGGCTTTAGGTT 1147

QY	985	GGTCTCAGCCCTGTGATGTTTGGAGCATAGTGTGCATCTTTATTTGAATATTTACCTTGTT	1044
DB	1148	GGTCTCAGCCCTGTGATGTTTGGAGCATAGTGTGCATCTTTATTTGAATATTTACCTTGTT	1207
QY	1045	TCAAGTCTTTTTCAGACTCATATAGTAAAGAGCCTGGCAATGATGGAAGCAATATATAG	1104
DB	1208	TCAAGTCTTTTTCAGACTCATATAGTAAAGAGCCTGGCAATGATGGAAGCAATATATATAG	1267
QY	1105	GACCCATACCAACACATGATGTTTCAAGAAACAGAAACGCAAGTATTTTCCACATAACC	1164
DB	1268	GACCCATACCAACACATGATGTTTCAAGAAACAGAAACGCAAGTATTTTCCACATAACC	1327
QY	1165	AGCTAGATTTGGGATGAAACACAGTCTCTGCTGTAGATATGTTAGGAGCGCTGCAAAACCGT	1224
DB	1328	AGCTAGATTTGGGATGAAACACAGTCTCTGCTGTAGATATGTTAGGAGCGCTGCAAAACCGT	1387
QY	1225	TGAAGGAATTTATGCTTTTGTATGATGAGAAACATGAGAACTGTTGACCTGTTGCA	1284
DB	1388	TGAAGGAATTTATGCTTTTGTATGATGAGAAACATGAGAACTGTTGACCTGTTGCA	1447
QY	1285	GAATGTTTGAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGAGCATCCTT	1344
DB	1448	GAATGTTTGAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGAGCATCCTT	1507
QY	1345	TCTTTGACTTTTAAAAAGAAATGAAATGGAATCAGTGTCTTATATATCTTCT	1404
DB	1508	TCTTTGACTTTTAAAAAGAAATGAAATGGAATCAGTGTCTTATATATCTTCT	1567
QY	1405	AGAAGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTTAATTTTGTAAACATT	1464
DB	1568	AGAAGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTTAATTTTGTAAACATT	1627
QY	1465	AAATTTATTTTGTACAGTTTAAAGTGAATATTTGTTTGTATCAATAGCATAATTAAC	1524
DB	1628	AAATTTATTTTGTACAGTTTAAAGTGAATATTTGTTTGTATCAATAGCATAATTAAC	1687
QY	1525	TTGTTAAGCAAGTATGCTTTGATAATGCAATAGAAAAATTTAAATTTTCTTTT	1584
DB	1688	TTGTTAAGCAAGTATGCTTTGATAATGCAATAGAAAAATTTAAATTTTCTTTT	1747
QY	1585	GAATTTACATTTTAAATACCTTTGAATATATCTTTGTGTCAGTGAATATGATTG	1644
DB	1748	GAATTTACATTTTAAATACCTTTGAATATATCTTTGTGTCAGTGAATATGATTG	1807
QY	1645	ATCTTGCTTTTGTATAGGAGTCACTCTGAGTGAATTTTGTAGTAAAGGAAT	1704
DB	1808	ATCTTGCTTTTGTATAGGAGTCACTCTGAGTGAATTTTGTAGTAAAGGAAT	1867
QY	1705	CTTGACTATTTTATATCTTTAAAGGAATTTCTTTTATATATCTTCAAAATTTAGAACTTAAC	1764
DB	1868	CTTGACTATTTTATATCTTTAAAGGAATTTCTTTTATATATCTTCAAAATTTAGAACTTAAC	1927
QY	1765	TTTAAAGTTTCTTCTGTAATTTGTTGAAGGGTGAATTTATTAACCTAGATAAGCA	1824
DB	1928	TTTAAAGTTTCTTCTGTAATTTGTTGAAGGGTGAATTTATTAACCTAGATAAGCA	1987
QY	1825	GCTACTAGAAACCAAACTCAGAAATCTTTTACTGTTAGAAATTTCTTAAATTTTAAAGT	1884
DB	2047	GCTACTAGAAACCAAACTCAGAAATCTTTTACTGTTAGAAATTTCTTAAATTTTAAAGT	2047
QY	1885	TTGATTTCTTTTTCATTTGGGTGATGTCAGGGTGAATAACAGACATTTATTTAAAGT	1944
DB	2048	TTGATTTCTTTTTCATTTGGGTGATGTCAGGGTGAATAACAGACATTTATTTAAAGT	2107
QY	1945	CAGTTTGTCCATTGTGACAGTTTGTTTTAAATAAACACATACACATTTTATTAAGT	2004
DB	2108	CAGTTTGTCCATTGTGACAGTTTGTTTTAAATAAACACATACACATTTTATTAAGT	2167
QY	2005	AAATCTTAAGTGAAGTCTGAAATGGAATTTTCCAAAGTATGTTTGTGAGTGCAC	2064
DB	2168	AAATCTTAAGTGAAGTCTGAAATGGAATTTTCCAAAGTATGTTTGTGAGTGCAC	2227
QY	2065	AGATATAAAATAGAAATTTCTGATGAGAGTTTTCAGTTTATTAATACCAAGTCTTTAGGAG	2124

DB	2228	AGATATAAAATAGAAATTTCTGATGAGAGGTTTTCAGTTTATTAATACCAAGTCTTTAGGAG	2287
QY	2125	TCTTAACATTTGGCAGCATCTGTTTATCAAAATGACATATAAAGCTTATAAGAAATT	2184
DB	2288	TCTTAACATTTGGCAGCATCTGTTTATCAAAATGACATATAAAGCTTATAAGAAATT	2347
QY	2185	AAAGTTTATTAATTTAGGCAATTTATGCTGTGATAATTTTACGGGAGAAAGAGATTGA	2244
DB	2348	AAAGTTTATTAATTTAGGCAATTTATGCTGTGATAATTTTACGGGAGAAAGAGATTGA	2407
QY	2245	TTGAAAGCAGTTTGGGAGAAAGTCTGCTGAAATTTCCAGAAATTTAATTTGATTGTTA	2304
DB	2408	TTGAAAGCAGTTTGGGAGAAAGTCTGCTGAAATTTCCAGAAATTTAATTTGATTGTTA	2467
QY	2305	CATAAACTTTTTCACCTTCA	2323
DB	2468	CATAAACTTTTTCACCTTCA	2486
RESULT 8			
AF294429			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
CDS			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1638; Conservative			
69.3%; Score 1632; DB 9; Length 1814;			
Pred. No. 3e-290;			
0; Mismatches 11; Indels 0; Gaps 0;			

```
QY 85 GTCAATTATTAGAACAGGTCCTTGAATGACGAGATTATCGGACCGGAGATACGTTG 144
Db 161 GTCAATTATTAGAACAGGTCCTTGAATGACGAGATTATCGGACCGGAGATACGTTG 220
QY 145 ACGAATACAGGAATGACATCTACTGTGTAAGGATATGTTCTTAGACATTATCACAGAGACATTG 204
Db 221 ACGAATAACGGAATGACATCTACTGTGTAAGGATATGTTCTTAGACATTATCACAGAGACATTG 280
QY 205 AAAGCGGTATCGAATCCACTGCACTGCACTTAATCTTCACTCCGAGCAGGAGAAAGCACTCTTA 264
Db 281 AAAGCGGTATCGAATCCACTGCACTGCACTTAATCTTCACTCCGAGCAGGAGAAAGCACTCTTA 340
QY 265 AAAGGAAGCGCAATAGACATCTGTTCAAGTCATCATCTGCTCAAGAGTGGAGACG 324
Db 341 AAAGGAAGCGCAATAGACATCTGTTCAAGTCATCATCTGCTCAAGAGTGGAGACG 400
QY 325 AAAGATCCAGAGATAGAGGATGATGAGGAGGTCACCTGATCTGCTCAAGAGTGGAGACG 384
Db 401 AAAGATCCAGGATAGAGGATGATGAGGAGGTCACCTGATCTGCTCAAGAGTGGAGACG 460
QY 385 TTCTAAGAGCAAGATATGAAATCGTGGACACTTTTGGGTGAAAGGACCTTTGGCAAGTTG 444
Db 461 TTCTAAGAGCAAGATATGAAATCGTGGACACTTTTGGGTGAAAGGACCTTTGGCAAGTTG 520
QY 445 TAGAGTCATTGATCATGCGCATGCGATGCGATGCTAGCAGTGAAGATCGTAAAGATG 504
Db 521 TAGAGTCATTGATCATGCGCATGCGATGCGATGCTAGCAGTGAAGATCGTAAAGATG 580
QY 505 TAGGCGGTACCGTCAAGCAGCTCGTTGAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTA 564
Db 581 TAGGCGGTACCGTCAAGCAGCTCGTTGAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTA 640
QY 565 CTGATCCCAATAGTGTCTCCGATGTCGAGATGCTAGAAATGTTGATCATCATGTC 624
Db 641 CTGATCCCAATAGTGTCTCCGATGTCGAGATGCTAGAAATGTTGATCATCATGTC 700
QY 625 ATGTTGTATTGTTGTTGAACTACTGCGGACTTAGTACTTACATTTCAATTAAGAAACA 684
Db 701 ATGTTGTATTGTTGTTGAACTACTGCGGACTTAGTACTTACATTTCAATTAAGAAACA 760
QY 685 GCTTTCTGCCATTTCAAAATTCGACATCCAGCAGATGCGGTATCAGATCTGCCAGTCAA 744
Db 761 GCTTTCTGCCATTTCAAAATTCGACATCCAGCAGATGCGGTATCAGATCTGCCAGTCAA 820
QY 745 TAAATTTTTTACATCATATAATTAATTAACCCATACAGATCTGAAAGCTGAAATATTTTGT 804
Db 821 TAAATTTTTTACATCATATAATTAATTAACCCATACAGATCTGAAAGCTGAAATATTTTGT 880
QY 805 TTGTGAAGTCTGATGATGATCAATATTAATTTCAAAATGAAGCTGATGAACGACAC 864
Db 881 TTGTGAAGTCTGATGATGATCAATATTAATTTCAAAATGAAGCTGATGAACGACAC 940
QY 865 TGAAAAACACAGATATCAAGTTGTTGACTTTGGAAGTGCACGATATCATGATGAACATC 924
Db 941 TGAAAAACACAGATATCAAGTTGTTGACTTTGGAAGTGCACGATATCATGATGAACATC 1000
QY 925 ACAGTACTTTGTTGTTCTACCGGCACTACAGAGCTCCGAGGTCAATTTTGGCTTTAGGTT 984
Db 1001 ACAGTACTTTGTTGTTCTACCGGCACTACAGAGCTCCGAGGTCAATTTTGGCTTTAGGTT 1060
QY 985 GGTCTCAGCTTTGATGTTGAGACATAGTTGATTTCTTATTTGAATATTACCTTGGTT 1044
Db 1061 GGTCTCAGCTTTGATGTTGAGACATAGTTGATTTCTTATTTGAATATTACCTTGGTT 1120
QY 1045 TCACAGTCTTCAGACTCATGATTAAGAGCAGCTGCGATGATGAAGCAATATTAG 1104
Db 1121 TCACAGTCTTCAGACTCATGATTAAGAGCAGCTGCGATGATGAAGCAATATTAG 1180
QY 1105 GACCCATACCAACACATGATTTCAGAAAAACGAAAGATTTTCCACCATACAC 1164
Db 1181 GACCCATACCAACACATGATTTCAGAAAAACGAAAGATTTTCCACCATACAC 1240
QY 1165 AGCTAGATTGGGATGAACACAGTTCTGCTGCTGCTGATATGTTAGGAGACGCTGCAACCGT 1224
```

```
Db 1241 AGCTAGATTGGGATGAACACAGTTCTGCTGCTGATGATATGTTAGGAGACGCTGCAACCGT 1300
QY 1225 TGAAGGAATTTATGCTTTGTCATGATGAAGAACATGAGAAACATGTTTGACCTGGTTCGAA 1284
Db 1301 TGAAGGAATTTATGCTTTGTCATGATGAAGAACATGAGAAACATGTTTGACCTGGTTCGAA 1360
QY 1285 GAATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGCAGCATCCTT 1344
Db 1361 GAATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGCAGCATCCTT 1420
QY 1345 TCTTTGACTTATTAAGAAAGAAATGAATGGAATCAGTGGTCTTACTATATACCTTCT 1404
Db 1421 TCTTTGACTTATTAAGAAAGAAATGAATGGAATCAGTGGTCTTACTATATACCTTCT 1480
QY 1405 AGAAGAGATTACTTAAGACTGTCAGTCAACTAAACATTTCTTAATTTTGTAAACATT 1464
Db 1481 AGAAGAGATTACTTAAGACTGTCAGTCAACTAAACATTTCTTAATTTTGTAAACATT 1540
QY 1465 AAATTAATTTTGTGTAAGTGAATATTTGATATGTTTGTATCAATAGCATATTAATTAAC 1524
Db 1541 AAATTAATTTTGTGTAAGTGAATATTTGATATGTTTGTATCAATAGCATATTAATTAAC 1600
QY 1525 TTGTTAAGCAAGTATGCTTTGATAATGCAATAGAAATTAATAATTTTCTTTT 1584
Db 1601 TTGTTAAGCAAGTATGCTTTGATAATGCAATAGAAATTAATAATTTTCTTTT 1660
QY 1585 GAAATTAACATTTTAAATACCTTTGAAATATCCTTTGTCAGTGAATAATGTGATTG 1644
Db 1661 GAAATTAACATTTTAAACATCCTTTGAAATATCCTTTGTCAGTGAATAATGTGATTG 1720
QY 1645 ATCTGCTTTTGTACATGAGGTCACCTCTGAAAGTGAATTTTGTGATGAAGAAAT 1704
Db 1721 ATCTGCTTTTGTACATGAGGTCACCTCTGAAAGTGAATTTTGTGATGAAGAAAT 1780
QY 1705 CTTGACTACTTTTATTTCTTAAAGGAATA 1733
Db 1781 CTTGACTACTTTTAAAGGAATA 1809

RESULT 9
LOCUS AX961896 1792 bp DNA linear PAT 14-JAN-2004
DEFINITION Sequence 107 from Patent WO03104277.
ACCESSION AX961896
VERSION AX961896.1 GI:40881306
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 Sugahara T., Matsuda A., Honda G., Muramatsu S. and Ishizawa K.
AUTHORS Stat6 activation gene
TITLE Patent: WO 03104277-A 107 18-DEC-2003;
JOURNAL Asahi Kasei Kabushiki Kaisha (JP)
FEATURES
source
1. 1792
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1. 1446
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAF06457.1"
/db_xref="GI:40881307"
/translation="MRHSKRTHCPDWDSRESNGHESYRSGHKKRKRSHSSSTQENRHCK
PHQFKESSDCHYLEARSLSNRDYYRNDYCEGYVPRHYRHRDIESGYRHS
KSSVRSRSPKRRKRHSQSHSRKRSRSEDDEGLHCOSGVLRYE
IVDTLGBGAFGVKVEICDHGMDGMHVAKIVKNGRYREARSIVQLWHLNSDPS
VFRQVQLEWPFHHGHVCIPELGLSTYDFIKENSLPFOIDHROMAYOICOSINF
LHNKLTDLKPENILFVKSQDYVKKNSKMRERILKNVDIKVDFGSATYDDEHH
STLVSTRHYRAPEVILALGWSQPCDVMWSIGCILIEYLIGLFTVFQTHDSKEHLAMMERI
```


LGPIPOHMIQKTRKRYFHHNQLDWDHSHSAGRYVRRRCKPLKEFMLCHDEHEKLPD
LVRMLEYDPTQRIITLDEALQHPFDLLKKK"

ORIGIN	Query Match	Best Local Similarity	Matches 1630; Conservative	Score 1629.4; DB 6; Length 1792;	Pred. No. 9.1e-290;	Mismatches 2;	Indels 0;	Gaps 0;
QY	85	GTCAATATTAGAGCAGGTCCTTGAATGAGCGAGATTATCGGGACCGGAGATACGTTG	144					
Db	161	GTCAATATTAGAGCAGGTCCTTGAATGAGCGAGATTATCGGGACCGGAGATACGTTG	220					
QY	145	ACGAATACAGGAATGACTACTGTGAAGGATATGTTCCCTAGACATTTATCACAGACATTG	204					
Db	221	ACGAATACAGGAATGACTACTGTGAAGGATATGTTCCCTAGACATTTATCACAGACATTG	280					
QY	205	AAAGGGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGCGAGGAGAGCGATCCTA	264					
Db	281	AAAGGGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGCGAGGAGAGCGATCCTA	340					
QY	265	AAAGGAGCGCAATAGACACTGTTCAAGTCATCACTGACGTTGCAAGAGCCACCGAAGGA	324					
Db	341	AAAGGAGCGCAATAGACACTGTTCAAGTCATCACTGACGTTGCAAGAGCCACCGAAGGA	400					
QY	325	AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAAGTCGAGACG	384					
Db	401	AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAAGTCGAGACG	460					
QY	395	TTCTAAGACGAGATATGAATCTGGACACTTTGGGTGAAGGAGCCCTTTGGCAAGTTG	444					
Db	461	TTCTAAGACGAGATATGAATCTGGACACTTTGGGTGAAGGAGCCCTTTGGCAAGTTG	520					
QY	445	TAGAGTGCAATGATCGCATGATGCGATGATGAGTACGATGCAAAATCGTAAAGAAATG	504					
Db	521	TAGAGTGCAATGATCGCATGATGCGATGATGAGTACGATGCAAAATCGTAAAGAAATG	580					
QY	505	TAGGCGGTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCACTTAAATAGTA	564					
Db	581	TAGGCGGTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCACTTAAATAGTA	640					
QY	565	CTGATCCCAATAGTCTCTCCGATGTCGAGATGCTAGATGTTTGAATCATCATCATGTC	624					
Db	641	CTGATCCCAATAGTCTCTCCGATGTCGAGATGCTAGATGTTTGAATCATCATCATGTC	700					
QY	625	ATGTTTGTATTGTGTTTCAACTACTGGGACTTAGTACTTACGATTTCAATTAAGAAACA	684					
Db	701	ATGTTTGTATTGTGTTTCAACTACTGGGACTTAGTACTTACGATTTCAATTAAGAAACA	760					
QY	685	GCTTCTGCCATTTCAAAATGACCAATGACGAGATGCGGTATCAGATCTGCCAGTCAA	744					
Db	761	GCTTCTGCCATTTCAAAATGACCAATGACGAGATGCGGTATCAGATCTGCCAGTCAA	820					
QY	745	TAAATTTTACATCATTAATAATTAACCCATACAGATCTGAGCCTGAAAATTTTGT	804					
Db	821	TAAATTTTACATCATTAATAATTAACCCATACAGATCTGAGCCTGAAAATTTTGT	880					
QY	805	TTGTGAAGTCTGACTATGTAGTCAATATAATTTCTAAATGAAACGTCATGAACGCAC	864					
Db	881	TTGTGAAGTCTGACTATGTAGTCAATATAATTTCTAAATGAAACGTCATGAACGCAC	940					
QY	865	TGAAAAACACAGATATCAAAAGTTGTTGACTTTGGAGGTGCAACGTCATGATGAACATC	924					
Db	941	TGAAAAACACAGATATCAAAAGTTGTTGACTTTGGAGGTGCAACGTCATGATGAACATC	1000					
QY	925	ACAGTACTTGGTGTCTACCGGCACTACAGAGTCCCGAGTCAATTTTGGCTTTAGGTT	984					
Db	1001	ACAGTACTTGGTGTCTACCGGCACTACAGAGTCCCGAGTCAATTTTGGCTTTAGGTT	1060					
QY	985	GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCAATCTTATTGAATATTACCTTGGTT	1044					
Db	1061	GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCAATCTTATTGAATATTACCTTGGTT	1120					
QY	1045	TCACAGTCTTTCAGACTCATGATAGTAAAGAGCACCTGGCAATGATGAACGAATATTAG	1104					

Db	1121	TCACAGTCTTTCAGACTCATGATAGTAAAGAGCACCTGGCAATGATGGACGAATATTAG	1180					
QY	1105	GACCCATACCAACACATGATTTCAGAAAAACAAGAAAAACGAAGTATTTTTCACCAATAACC	1164					
Db	1181	GACCCATACCAACACATGATTTCAGAAAAACAAGAAAAACGAAGTATTTTTCACCAATAACC	1240					
QY	1165	AGCTAGATTGGGATGAACACAGTTCCTGCTGGTATATGTTAGGAGACGCTGCAAAACCGT	1224					
Db	1241	AGCTAGATTGGGATGAACACAGTTCCTGCTGGTATATGTTAGGAGACGCTGCAAAACCGT	1300					
QY	1225	TGAAGGAATTTATGCTTTGTCATGATGAAGAAACATGAGAACTGTTTGACCTGGTCGAA	1284					
Db	1301	TGAAGGAATTTATGCTTTGTCATGATGAAGAAACATGAGAACTGTTTGACCTGGTCGAA	1360					
QY	1285	GAATGTTAGAAATATGATCCAACTCAAAGAAATTAACCTTGGATGAAGCAATTCGACGATCCTT	1344					
Db	1361	GAATGTTAGAAATATGATCCAACTCAAAGAAATTAACCTTGGATGAAGCAATTCGACGATCCTT	1420					
QY	1345	TCCTTGAATTTTAAAGAAATGAATGGGAATCAGTGGTCTTACTATATATCTTCTCT	1404					
Db	1421	TCCTTGAATTTTAAAGAAATGAATGGGAATCAGTGGTCTTACTATATATCTTCTCT	1480					
QY	1405	AGAAGAGATTACTTAAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATT	1464					
Db	1481	AGAAGAGATTACTTAAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATT	1540					
QY	1465	AAATATTTTGTACAGTTAAGTGTAAATATTTGATGTTTGTATCAATCAATGAATAATTAAC	1524					
Db	1541	AAATATTTTGTACAGTTAAGTGTAAATATTTGATGTTTGTATCAATCAATGAATAATTAAC	1600					
QY	1525	TTGTTAAGCAAGTATGGTCTTGAATAATGCAATTAAGAAAAATAAATTTTCTTTT	1584					
Db	1601	TTGTTAAGCAAGTATGGTCTTGAATAATGCAATTAAGAAAAATAAATTTTCTTTT	1660					
QY	1585	GAATTTACCAATTTTAAATACCTTTGAAATATCCTTTGTCGCCAGTCATAAATGTGATTG	1644					
Db	1661	GAATTTACCAATTTTAAATACCTTTGAAATATCCTTTGTCGCCAGTCATAAATGTGATTG	1720					
QY	1645	ATCTGCTTTTGTACATGGAGGTCACTCTGAAGTGAATTTTTTTTTCAGTAAAGGAAAT	1704					
Db	1721	ATCTGCTTTTGTACATGGAGGTCACTCTGAAGTGAATTTTTTTTTCAGTAAAGGAAAT	1780					
QY	1705	CTTGACTACTTT 1716						
Db	1781	CTTGACTACTTT 1792						

RESULT 10
BC063116
LOCUS
DEFINITION Homo sapiens CDC-like kinase 4, mRNA (cdna clone IMAGE:30324121), partial cds.
ACCESSION BC063116
VERSION BC063116.1 GI:38970005
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 3492)
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshyuki, S., Abramson, R.D., Mullighy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahney, J., Helton, E., Kerteman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grinwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzyzaniak, M.I., Skalska, U., Smalls, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1' to 3492)

Strausberg, R.

Direct Submission

Submitted (02-DEC-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits

cDNA library preparation: CLONTECH Laboratories, Inc.

cDNA library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

info@cgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letlicia Hsiao, Martin Krzyzinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jacqueline George Yang, Scott Zuyderdyk, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 51 Row: a Column: 9
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA i: 10190705.

FEATURES
source

```

1. 3492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30324121"
/tissue_type="Skin and m
pia matter, choroid plexu
/clone_lib="NIH_MGC_186"
/lab_host="DH10B"
note="Vector: pDNR-LIB"

```

ORIGIN

```
Query Match          56.0%; Score 1318.4; DB 9; Length 3492;
Best Local Similarity 99.2%; Pred. NO. 1.3e-232;
Matches 1325; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

Qy	398	ATATGAAATCGTGGACACTTTTGGGTGAAGGAGCCTTTGGCGAAGTTGTAGTGCATTGA	457
Db	2149	AGATGAAATCGTGGACACTTTTGGGTGAAGGAGCCTTTGGCGAAGTTGTAGTGCATTGA	2208

QY	458	TCA	TGG	CAT	TGG	ATG	GG	CAT	TGC	ATG	CAG	TGA	AAAA	ATCG	TG	TAG	GGC	CGT	TAC	CG	517
Db	2209	TCA	TGG <td>CAT</td> <td>TGG <td>ATG</td> <td>GG</td> <td>CAT</td> <td>TGC <td>ATG</td> <td>CAG</td> <td>TGA <td>AAAA</td> <td>ATCG <td>TG</td> <td>TAG</td> <td>GGC <td>CGT <td>TAC <td>CG</td> <td>2268</td> </td></td></td></td></td></td></td>	CAT	TGG <td>ATG</td> <td>GG</td> <td>CAT</td> <td>TGC <td>ATG</td> <td>CAG</td> <td>TGA <td>AAAA</td> <td>ATCG <td>TG</td> <td>TAG</td> <td>GGC <td>CGT <td>TAC <td>CG</td> <td>2268</td> </td></td></td></td></td></td>	ATG	GG	CAT	TGC <td>ATG</td> <td>CAG</td> <td>TGA <td>AAAA</td> <td>ATCG <td>TG</td> <td>TAG</td> <td>GGC <td>CGT <td>TAC <td>CG</td> <td>2268</td> </td></td></td></td></td>	ATG	CAG	TGA <td>AAAA</td> <td>ATCG <td>TG</td> <td>TAG</td> <td>GGC <td>CGT <td>TAC <td>CG</td> <td>2268</td> </td></td></td></td>	AAAA	ATCG <td>TG</td> <td>TAG</td> <td>GGC <td>CGT <td>TAC <td>CG</td> <td>2268</td> </td></td></td>	TG	TAG	GGC <td>CGT <td>TAC <td>CG</td> <td>2268</td> </td></td>	CGT <td>TAC <td>CG</td> <td>2268</td> </td>	TAC <td>CG</td> <td>2268</td>	CG	2268

QY	518	TGAAGCAGCTCGTTCAGAAATCCAA	GATTAGAGCACTTAAATAGTACTGAT	CCCCAATAG	577
Db	2269	TGAAGCAGCTCGTTCAGAAATCCAA <th>GATTAGAGCACTTAAATAGTACTGAT</th> <th>CCCCAATAG</th> <th>2328</th>	GATTAGAGCACTTAAATAGTACTGAT	CCCCAATAG	2328

Oy		578	TGTCCTCCGATGTGTTCCAGATGCTAGAAATGGTTTTGATCATCATGGTGCATGTTTGTAATTGT	637
Dd		2329	TGTCCTCCGATGTGTTCCAGATGCTAGAAATGGTTTTGATCATCATGGTGCATGTTTGTAATTGT	2388

QY	638	GTGTTGAACCTACTCTGGGACCTTAGTACTTACGATTTTCATTAAGAAACACGCTTTCCTGCCATT	6397
Db	2389	GTGTTGAACCTACTCTGGGACCTTAGTACTTACGATTTTCATTAAGAAACACGCTTTCCTGCCATT	2448
QY	698	TCAAATTTGACCAATCATCAGGCGAGATGCGGTATCAGATCTGCGAGTCAATAAATTTTTTACA	757
Db	2449	TCAAATTTGACCAATCATCAGGCGAGATGCGGTATCAGATCTGCGAGTCAATAAATTTTTTACA	2508
QY	758	TCATAATAAATTAACCCATACAGATCTGAAGCCTGAAATATTTTGTGTGAGTCTCGA	817
Db	2509	TCATAATAAATTAACCCATACAGATCTGAAGCCTGAAATATTTTGTGTGAGTCTCGA	2568
QY	818	CTATGTAGTCAAAATATAATTTCTAAATGAAAAGTGAAGCGACACACTGAAAACACAGA	877
Db	2569	CTATGTAGTCAAAATATAATTTCTAAATGAAAAGTGAAGCGACACACTGAAAACACAGA	2628
QY	878	TATCAAAGTTGTTGACTTTTGGAGTGCACGTATGATGATGAACATCACAGTACTTTGGT	937
Db	2629	TATCAAAGTTGTTGACTTTTGGAGTGCACGTATGATGATGAACATCACAGTACTTTGGT	2688
QY	938	GTCTACCCGGCCTACTACAGAGCTCCGAGGTCAATTTTGGCTTTAGTTGGTCTCTCAGGCTTG	997
Db	2689	GTCTACCCGGCCTACTACAGAGCTCCGAGGTCAATTTTGGCTTTAGTTGGTCTCTCAGGCTTG	2748
QY	998	TGATGTTTGGAGCATAGGTTGCATTCCTATTGAATATTAACCTTGGTTTACAGTCTTTTCA	1057
Db	2749	TGATGTTTGGAGCATAGGTTGCATTCCTATTGAATATTAACCTTGGTTTACAGTCTTTTCA	2808
QY	1058	GACTCATGATAGTAAAGAGCCTCGCNAATGATGGAAAGCAATATTAGGACCCATACCACA	1117
Db	2809	GACTCATGATAGTAAAGAGCCTCGCNAATGATGGAAAGCAATATTAGGACCCATACCACA	2868
QY	1118	ACACATGATTCAGAAAAACAAGAAAAACCAAGTATTTTACCATAAACAGCTAGATTGGGA	1177
Db	2869	ACACATGATTCAGAAAAACAAGAAAAACCAAGTATTTTACCATAAACAGCTAGATTGGGA	2928
QY	1178	TGAACAAGTTCTGCTGGTAGATATGTTTAGGAGACGCTGCAAAACCGTTGAAGGAATTTAT	1237
Db	2929	TGAACAAGTTCTGCTGGTAGATATGTTTAGGAGACGCTGCAAAACCGTTGAAGGAATTTAT	2988
QY	1238	GCCTTTGTCATGATGAAGACATGAGAAAATCTGTTTGACCTGGTTGCGAAGATGTTAGAATA	1297
Db	2989	GCCTTTGTCATGATGAAGACATGAGAAAATCTGTTTGACCTGGTTGCGAAGATGTTAGAATA	3048
QY	1298	TGATCCAACTCAAAGAAATTTACCTTGGATGAAGCATTTGCAGCATCCCTTCTTTGACCTTAT	1357
Db	3049	TGATCCAACTCAAAGAAATTTACCTTGGATGAAGCATTTGCAGCATCCCTTCTTTGACCTTAT	3108
QY	1358	AAAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATCTTCTCTAGAGAGATTTACT	1417
Db	3109	AAAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATCTTCTCTAGAGAGATTTACT	3168
QY	1418	TAAGACTGTGTCAGTCAACTAAACATTTCTAAATATTTTGTGTAACATTAATTTTGTGA	1477
Db	3169	TAAGACTGTGTCAGTCAACTAAACATTTCTAAATATTTTGTGTAACATTAATTTTGTGA	3228
QY	1478	CAGTTAAAGTGTAATAATGATGTTTGTGTTGATCAATAGCATAAATTAATTTGTAAGCAAGT	1537
Db	3229	CAGTTAAAGTGTAATAATGATGTTTGTGTTGATCAATAGCATAAATTAATTTGTAAGCAAGT	3288
QY	1538	ATGGTCTTCGATAATGCATTTAGAAAAATTAATAATTAATTTTCTTTTGGAAATTTACCATTT	1597
Db	3289	ATGGTCTTCGATAATGCATTTAGAAAAATTAATAATTAATTTTCTTTTGGAAATTTACCATTT	3348
QY	1598	TTAAATACCTTTTGAATAATATCCTTTTGTGTCGAGTGAATAATGTAATGATCTTGCCCTTTTG	1657
Db	3349	TTAAATACCTTTTGAATAATATCCTTTTGTGTCGAGTGAATAATGTAATGATCTTGCCCTTTTG	3408
QY	1658	TACATGGAGGTCACTCTCAAGTGATTTTTTTTTTGGAGTAAAAAGGAATCTTGACTACTTTTA	1717
Db	3409	TACATGGAGGTCACTCTCAAGTGATTTTTTTTTTGGAGTAAAAAGGAATCTTGACTACTTTTA	3468
QY	1718	TATTTCTTAAAGGAATA	1733

Db 3469 AAAAAAAAAAAAAAAAAA 3484

|||||

RESULT 11
AY335726
LOCUS
DEFINITION
AY335726 1446 bp mRNA linear SYN 15-OCT-2003
Synthetic construct Homo sapiens CDC-like kinase 4 (CLK4) mRNA,
partial cds.

ACCESSION
AY335726.1 GI:33304090
VERSION
FLI_CDNA.
KEYWORDS
synthetic construct
SOURCE
synthetic construct
ORGANISM
other sequences; artificial sequences.
REFERENCE
1 (bases 1 to 1446)
Park,J., Rolfs,A., Hu,Y., Shen,B., Vannberg,F., Moreira,D.,
Kelley,T., Zuo,D., Raphael,J., Baqui,M., Jepson,D., Harlow,E.,
LaBaer,J. and Brizuela,L.
AUTHORS
Cloning of human full-length CDS FLEXGene kinases in
recombinational vector system
TITLE
Unpublished
JOURNAL
2 (bases 1 to 1446)
Park,J., Rolfs,A., Hu,Y., Shen,B., Vannberg,F., Moreira,D.,
Kelley,T., Zuo,D., Raphael,J., Baqui,M., Jepson,D., Harlow,E.,
LaBaer,J. and Brizuela,L.
REFERENCE
Direct Submission
AUTHORS
Submitted (02-JUL-2003) Biological Chemistry and Molecular
JOURNAL
Pharmacology, Harvard Institute of Proteomics, 320 Charles St.,
Cambridge, MA 02141-2023, USA
COMMENT
This CDS clone is a part of a collection of human full-length
expression clones generated by Harvard Institute of Proteomics.
Each CDS has been cloned without stop-codon (to allow fusion with
C-terminal tag). The CDS has been directionally cloned using BD
In-Fusion (TW) cloning system between the Sali and HindIII sites of
the pDNR-Dual vector. Additional sequences in the clone: 'ACC'
after Sali site and before 'ATG' to provide Kozak consensus
sequence; 'GG' after last codon and before HindIII site to maintain
reading frame.

FEATURES
source
Location/Qualifiers
1..1446
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/clone="FLH000244.01L"
/tissue_type="placenta and brain"
/clone_lib="First strand cDNA from placenta and brain"
/lab_host="DH5alpha T1 resistant"
/notes="Vector: pDNR-Dual"
1..>1446
/genes="CLK4"
1..>1446
/genes="CLK4"
/notes="Mutations: Stop->Leu"
/codon_start=1
/transl_table=11
/product="CDC-like kinase 4"
/protein_id="AAQ02553.1"
/db_xref="GI:33304091"
translation="MRSRKTHCFDWDRSRGHESYRGSHKKRKRSHSSTQBNRRCK
PHQFKESDCHYLLEARNEDRYDYNDYCEGYVPRHYRDIIESGYRIHCS
KSVRRGSRPKRKNRHCSSHOSKSHRRKRSRSTEDDEGHEGLICQSDGLVDRP
IVDTLGGAGKQVECTDHGMCGMHVAKIVKNGYRRAARSEIOLVLEHLNSTDRN
VFCVQMLEWFDHHGHVCIVPELLGLSTYDFIKENSFLPQIDHIIQIMAYQICQSI
LHNKTHLTKDPEVILFVKSDYVKNYSKQDKERTLKNTDIKVVDGSGATVDDSHH
STLVTSYRRAPEVILALGWSQPCDWSIGCILIEYLGTFVTFQSDKSEHLAMMERI
LGSPIDPMIOIKTRKRYFHHNQLDWDEHSAGSYRRRCCKPLKFEMLCHDEEKLFL
LVRRMLEYDPTQRTITLDEALQHPFDLLKKKL"

gene
CDS

ORIGIN
Query Match 54.5%; Score 1284; DB 12; Length 1446;
Best Local Similarity 100.0%; Pred. No. 3.2e-226;
Matches 1284; Conservative 0; Mismatches 0; Indels 0; Caps 0;

681 TAGGCCGTTACCGGAGGAGCAGTCGTTCTGAAATCCAAAGTATTTGGAGCACTTGAACAGCA 740
 565 CTGATCCCAATAGTGTCTTCGAGTGTGTCGAGTGTAGATGTTTGTATCATCATGTC 624
 741 CTGACCCCAACAGTGTCTTCGAGTGTGTCGAGTGTAGATGTTTGTATCATCATGTC 800
 625 ATGTTTGTATTTGTGTTGAACTTACTGGGACTTAGTACTTACGATTTTCAATTAAGAAACA 684
 801 ATGTTTGTATTTGTGTTGAGCTGCTGGGACTTAGTACTTACGATTTTCAATTAAGAAACA 860
 685 GCTTTCTGCCATTTCAAAATGACACATCAGCAGATGCGGTATCAGATCTGCCAGTCAA 744
 861 GTTTTCTGCCATTTCAAAATGATCAGTGTGCTGGGACTTAGTACTTACGATTTTCAATTAAGAAACA 920
 745 TAAATTTTTCATCATATAAATTAACCCATACAGATCTGAGCTGCAAGCTGCAAAATTTTGT 804
 921 TAAATTTTTCATCATATAAATTAACACACAGGACTTAAACCTGCAAAATTTTGTATTTAT 980
 805 TTGTAAGTCTGACTATGCTAGTCAAAATATAATTTCAAAATGAAACGTCATGAAACGACAC 864
 981 TTGTAAGTCTGACTATGCTAGTCAAAATATAATTTCAAAATGAAACGTCATGAAACGACAC 1040
 865 TGAATAACACAGATATCAAAATGTTGTGACTTTTGGAAAGTGCACGATGATGATGAACATC 924
 1041 TGAATAACACAGATATCAAAATGTTGTGACTTTTGGAAAGTGCACATATGACGACGACATC 1100
 925 ACAGTACTTTGGTGTCTACCCGGCACTACAGAGCTCCGAGGTGCAATTTTGGCTTTAGGTT 984
 1101 ATAGTACTTTGGTGTCTACCAAGGCACTACAGGCTCCAGAGGTGCAATTTTGGCTTTAGGTT 1160
 985 GGTCTCAGCCTTTGATGTTTGGAGCATAGTGTGCAATTTTATTAATTAATTAATTAATTAAT 1044
 1161 GGTCTCAGCCTTTGATGTTTGGAGCATAGTGTGCAATTTTATTAATTAATTAATTAATTAAT 1220
 1045 TCACAGTCTTCAGACTCATGATAGTAAAGACACCTGGCAATGATGAAAGATTAATTAATTAAT 1104
 1221 TCACAGTCTTCAGACTCATGATAGTAAAGACACCTGGCAATGATGAAAGATTAATTAATTAAT 1280
 1105 GACCATACCAACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1164
 1281 GACCATACCAACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1340
 1165 AGTAGATTTGGATGAAACACAGTCTTCGCTGCTGATGATGATGATGATGATGATGATGATGAT 1224
 1341 AGTAGATTTGGATGAAACACAGTCTTCGCTGCTGATGATGATGATGATGATGATGATGATGAT 1400
 1225 TGAAGGAATTTATGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1284
 1401 TAAAGGAATTTATGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1460
 1285 GAAAGTGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1344
 1461 GAAAGTGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1520
 1345 TCTTTGACTTTTAAAGGAATGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 1404
 1521 TCTTTGACTTTTAAAGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 1579
 1405 AGAAGAGATTTACTTAAGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1464
 1580 GGAGCAGTACTTCCAGAGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1639
 1465 AATTTATTTTGTACAGTGTAAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1524
 1640 AATTTATTTTGTACAGTGTAAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1699
 1525 TTGTTAAGCAAGTATGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1583
 1700 TTGTTAAGTATGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1759
 1584 TGAATTTACCAATTTTAAATACCTTTGAAATATCCTTTGTTGTCAGTGTAAATGATGATGATGAT 1642
 1760 ATGTTAATTTGCGGCTTTAAAGGCTTTTAAAGGCTTTTAAAGGCTTTTAAAGGCTTTTAAAGGCT 1819

QY 1643 TGATCTTCCCTTTTGTATCATGAGGTACCTCTCTGAAGTGAATTTTGTAGTAAAGGAA 1702
 Db 1820 TGGTCCCATCTTTTGTATCATGAGGTGACCTCTGAAGTGAATTTTGTAGTAAAGGAA 1879
 QY 1703 A 1703
 Db 1880 A 1880
 RESULT 13
 AX961894 1881 bp DNA linear PAT 14-JAN-2004
 LOCUS Sequence 105 from Patent WO03104277.
 DEFINITION
 AX961894
 ACCESSION
 VERSION AX961894.1 GI:40881304
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 1 Sugahara,T., Matsuda,A., Honda,G., Muramatsu,S. and Ishizawa,K.
 Stat6 activation gene
 Patent: WO 03104277-A 105 18-DEC-2003;
 Asahi Kasei Kabushiki Kaisha (JP)
 JOURNAL
 FEATURES
 Location/Qualifiers
 source 1..1881
 /organism="Mus musculus"
 /mol_type="unassigned DNA"
 /db_xref="taxon:10090"
 90..1535
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAP06456.1"
 /db_xref="GI:40881305"
 /translation="MRHSKRTHCPDWDPSRESWGESYSGSHKRRKRSHSSTQENRHCK
 PHGFQSDCHYLEARCLNEDYRDYRIDEYNDYCEGYVPRHYHRDVEYTYRHCS
 KSSVRSRSPKRRKPCASHQSHSRKRSIEDDEGHLICQSGDVLRYRVE
 IVDTLGSGAGKVEECIDHMGDLHVAKI VKNVGRVREAAARSETQVLEHLNSTPNS
 VFRQVQLEWDFDHGHVCIPEYGLSLSTYDFIKENSLPFQIDHROMAYOICQSLNF
 LHNKLTHDLPENILFVSDYVYVKNKSKMRDERILKNVDIKVDFGSGATYDDEHH
 STLYSTRHYRAPEVILALGWSQPCDWSIGCILLEYLGFVTFQTHDSKEHLAMMERI
 LGTPAPAKMIQKTRKRYFHNQLDDEHDSAGRYVRRRCRKLKSFMLCHDEHEKLF
 LVRRMLEYDPPARRITLDELQHPFDLLKRR"
 ORIGIN
 Query Match 54.3%; Score 1277.4; DB 6; Length 1881;
 Best Local Similarity 87.9%; Pred. No. 5.1e-225;
 Matches 1413; Conservative 0; Mismatches 192; Indels 3; Gaps 2;
 QY 85 GTCATTATTTAGAACGAGGTCTTTGAATGAGGAGATTTATCGGACCGGAGATACGTTG 144
 Db 250 GTCATTATTTAGAACGAGGTCTTTGAATGAGGAGATTTATCGGACCGGAGATACGTTG 309
 QY 145 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTATCACAGACATTTG 204
 Db 310 ATGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTATCACAGACATTTG 369
 QY 205 AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCTAGTCCGACGAGGAGGAGGAGTCTTA 264
 Db 370 AAAGCACTTACCGGATCCATTTGCAGTAAATCTCTAGTCCGAGGAGGAGGAGGAGGAGTCTTA 429
 QY 265 AAAGGAGCGCAATGACACACTGTTTCAAGTCTACGTACGTTCGAAGAGGAGGAGGAGGAGGAG 324
 Db 430 AGAGAAACGGTAATAGACCCCTGTGCAAGTCACTAGTCCGATTCGAGAGGAGGAGGAGGAGGAG 489
 QY 325 AAAGTCCAGGAGTATAGAGGATGATGAGGAGGTACCTGATCTGTCAAAGTGGAGAGG 384
 Db 490 AAAGTCCAGGAGTATAGAGGATGATGAGGAGGTACCTGATCTGTCAAAGTGGAGAGG 549
 QY 385 TTCTAAGAGGAGATATGAATCTGGACACTTTGGGTGAAGGAGGCTTTTGGCAAGTTG 444

Db 550 TTCTAAGCAAGATATGAATCGTGGACACTTTAGGTGAAGGAGCCCTTTGGCAAGTTG 609
Qy 445 TAGAGTGCATTCATGATGGCATGGCATGCTAGTACAGTGAATAATCGTAAAAAATG 504
Db 610 TAGAGTGCATTCATGACGGCATGGCTTACATGTAGCAGTGAATAATCGTAAAAAATG 669
Qy 505 TAGGCGGTACCGTGAAGCAGCTGTTCAGAAATCCAAAGTATTAGAGCAGCTTAATAATGTA 564
Db 670 TAGGCGGTACCGGAGGAGCAGCTGTTCGAAATCCAAAGTATTGGAGCAGCTTTGAACGCA 729
Qy 565 CTGATCCCAATAGTGTCTTCGAGTGTCCAGATGCTAGAGTGTGTTGATCATCATGTGTC 624
Db 730 CTGACCCCAACAGTGTCTTCGAGTGTCCAGATGCTAGAGTGTGTTGATCATCATGTGTC 789
Qy 625 ATGTTTGTATTGTGTTTGAACCTACTGGGACTTAGTACTTACGATTTTCAATTAAGAAAAA 684
Db 790 ATGTTTGTATTGTGTTTGAACCTACTGGGACTTAGTACTTACGATTTTCAATTAAGAAAAA 849
Qy 685 GCTTCTGCCAATTCAAATGACACATCAGGAGATGGCGTATCAGATCTGCCAGTCAA 744
Db 850 GTTTTCTGCCAATTCAAATGATCACATCAGGCAAAATGGCTTATCAGATCTGCCAGTCTA 909
Qy 745 TAAATTTTTCATCATATAATTAACCCATACAGATCTCAAGCCTGAATAATTTTGT 804
Db 910 TAAATTTTTCATCATATAATTAACACACAGGACCTAAACCTTGAAAAATATTTTAT 969
Qy 805 TTGTGAAGTCTGACTATGACTCAAAATATAATTTCAAAATCAAACTGATCAAGCCACAC 864
Db 970 TTGTGAAGTCTGACTATGACTCAAAATATAATTTCAAAATCAAACTGATCAAGCCACAT 1029
Qy 865 TGAATAAACACAGATATCAAAAGTTGTGACTTTTGAAGTGCACGATGATGATGAACATC 924
Db 1030 TGAATAAACACAGATATCAAAAGTTGTGACTTTTGAAGTGCACATATCAGCAGCAACATC 1089
Qy 925 ACAGTACTTTTGGTCTACCGGCACATCAGAGCTCCGAGGTCAATTTTGGCTTTAGTT 984
Db 1090 ATAGTACTTTTGGTCTCACAAGGCACTACAGGCTCCAGAGGTCAATTTTGGCTCTAGTT 1149
Qy 985 GGTCTCAGCCTTTGTGATGTTTGGAGCATAGTGTGATTTTATTTGAATATTAACCTTTGTT 1044
Db 1150 GGTCTCAGCCTTTGTGATGTTTGGAGCATAGTGTGATTTTATTTGAAGTACTTGGTT 1209
Qy 1045 TCAGAGTCTTTTCACTCATGATAGTAAAGACACCTGGCAATGATGGAACGAATATPAG 1104
Db 1210 TCAGAGTCTTTTCACTCATGATAGTAAAGACACCTGGCAATGATGGAAGGAGTCTTAG 1269
Qy 1105 GACCCATACCAACACATGATTCAGAAACCAAGAAACGCAAGTATTTTCAACATAACC 1164
Db 1270 GACCCATCCAGCATATGATTCAGAAACCAAGAAACGCAAGTATTTTCAACATAACC 1329
Qy 1165 AGCTAGATTGGATGAACACAGTTCTGCTGTGTAGATATGTTAGGAGAGCGCTGCAACCGT 1224
Db 1330 AGCTAGATTGGATGAACACAGTTCTGCTGTGTAGATATGTTAGGAGAGCGCTGCAACCGT 1389
Qy 1225 TGAAGGAATTTATGCTTTGTCATGATGAAGAACATGAGAAATCTGTTGACCTGGTTGCA 1284
Db 1390 TAAAGGAATTTATGCTGTGTGTCATGACGAAGACATGAGAAAGCTGTTTCACTGGTTGCA 1449
Qy 1285 GAATGTTAGATATGATCCAACTCAAGAAATACCTTGGATGAAGCAATGAGCATTCCT 1344
Db 1450 GAATGTTGAGTATGATCCAGCAGGAAGGATCACCTTGGATGAAGCAATGAGCATTCCT 1509
Qy 1345 TCCTTGACTTATTAAGAAAGAAATGAAATGGAATCAGTGTCTTACTATATATCTTCT 1404
Db 1510 TCCTTGACTTATTAAGAAAGAAATGAAATGGAATCAGTGTCTTACTATATATCTTCT 1568
Qy 1405 AGAAGAGATTAATTAAGACTGTGTGCTCAGTCAACTAAACATTTCTAATATTTTGTAAACAT 1464
Db 1569 GGAGCAGTTACTTCCAGACTGTGTGCTCAGTCAACTAAACCTTCTAATATTTTGTAAACAT 1628
Qy 1465 AAATATTTTGTACAGTTAAGTGAATAATTTGTATGTTTGTATCAATAGCATTAATTAAC 1524
Db 1629 AAATATTTTGTACAGTTAAGTGAATAATTTGTATGTTTGTATCAATAGCATTAATTAAC 1688

Qy 1525 TTGTTAAGCAAGTATGCTTGTGATAATGCAATAGAAAAATTAATTAATTTTCTTTT- 1583
Db 1689 TTGTTAAGTATGCTTGTGATAATGAATGAATCAATCAAGAGTTTCAATGAATTTCTTTTG 1748
Qy 1584 -TGAATAATCACTTTTAAATACCTTTGAATATATCCTTTGTGTCAGATGAATAATGTCAT 1642
Db 1749 ATGTTAATGTCACCTTTTAAAGGCTTTTAGAATGCCCCCTTTGTCCAGTGAATAAGTGTAT 1808
Qy 1643 TGATCTTGGCCCTTTGTATACATGGAGTCACTCTCTCAAGTGAATTTTCTTTT 1690
Db 1809 TGGTCCCATCTTTTGTATCATGAAGTTCGACTCTCAAGTGAATTTTCTTTT 1856

RESULT 14
BD260616
LOCUS BD260616 1222 bp DNA linear PAT 17-JUL-2003
DEFINITION 49 human secreted proteins.
ACCESSION BD260616
VERSION BD260616.1 GI:33070386
KEYWORDS JP 2002539776-A/16.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1222)
AUTHORS Rosen,C.A., Ruben,S.M. and Komatsuoulis,G.
TITLE 49 human secreted proteins
JOURNAL Patent: JP 2002539776-A 16 26-NOV-2002;
HUMAN GENOME SCIENCES INC
COMMENT OS Homo sapiens (human)
PN JP 2002539776-A/16
PD 26-NOV-2002
PF 16-NAR-2000 JP 2000606614
PR 23-WAR-1999 US 60/126051,10-DEC-1999 US 60/169906 PI
CRAIG A ROSEN,STEVEN M RUBEN,GEORGE KOMATSUOULIS PC
C12N15/09,A61K38/36,A61K38/00,A61K49/00,A61P17/02,A61P25/14, PC
A61P25/16,
PC A61P25/18,A61P25/24,A61P25/28,A61P29/00,C07K14/47,C07K16/18,
PC C12N1/15,
PC C12N1/19,C12N1/21,C12N5/10,C12P21/02,C12Q1/02,C12Q1/68,G01N33/
PC 53,
PC G01N33/566,G01N33/68,C12N15/00,A61K37/02,C12N5/00 CC 49
human secreted proteins
FH Key Location/Qualifiers
FT source 1..1222
FT /organism='Homo sapiens (human)'.
FEATURES
source
Location/Qualifiers
1..1222
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 50.3%; Score 1185; DB 6; Length 1222;
Best Local Similarity 99.9%; Pred. No. 5.4e-208;
Matches 1196; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 517 GTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCACTTAATAGTACTGATCCCAATA 576
Db 9 GTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCACTTAATAGTACTGATCCCAATA 68
Qy 577 GTGCTTCCGATGTGTCAGATGCTAGAAATGGTTTGTATCATCATGGTGTGTTGTATTG 636
Db 69 GTGCTTCCGATGTGTCAGATGCTAGAAATGGTTTGTATCATCATGGTGTGTTGTATTG 128
Qy 637 TGTTTGAATCTACTCGGACTTAGTACTTACGATTTTCAATTAAGAAACAGCTTTCTGCCAT 696
Db 129 TGTTTGAATCTACTCGGACTTAGTACTTACGATTTTCAATTAAGAAACAGCTTTCTGCCAT 188
Qy 697 TTCAATTTGACCAATCAGGAGATGGGATATCAGATCTGCGAGTCAATAATTTTCTAC 756
Db 189 TTCAATTTGACCAATCAGGAGATGGGATATCAGATCTGCGAGTCAATAATTTTCTAC 248

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2005, 01:13:38 ; Search time 1244.4 Seconds

(without alignments)
11198.198 Million cell updates/sec

Title: US-10-801-671-1

Perfect score: 2354

Sequence: 1 gccagcgtgggtactttaa.....taaaaaaaaaaaaaaaaaaaac 2354

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04.*
1: geneseqn1980s.*
2: geneseqn1990s.*
3: geneseqn2000s.*
4: geneseqn2001as.*
5: geneseqn2001bs.*
6: geneseqn2002as.*
7: geneseqn2002bs.*
8: geneseqn2003as.*
9: geneseqn2003bs.*
10: geneseqn2003cs.*
11: geneseqn2003ds.*
12: geneseqn2004as.*
13: geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2354	100.0	2354	6 AAD32038	Aad32038 Human kin
2	2215.8	94.1	2497	4 AAF89402	Aaf89402 Human cel
3	2052.6	87.2	2446	2 AAX89852	Aax89852 Human pro
4	2048	87.0	2488	4 Aaf44675	Aaf44675 Novel pro
5	2048	87.0	2488	12 ADI29373	Adi29373 Human MAR
6	1632	69.3	1814	13 ADP24393	Adp24393 PRO poly
7	1623.4	69.2	1792	12 ADI26142	Adi26142 Human CDN
8	1300.6	55.3	1323	3 AAC81772	Aac81772 Human bea
9	1286	54.6	1446	12 ADO07809	Ado07809 Human pol
10	1277.4	54.3	1881	12 ADI26140	Adi26140 Human CDN
11	1185	50.3	1222	3 AAC59283	Aac59283 Human sec
12	1178.2	50.1	1456	11 ADI31373	Adi31373 Human CDN
13	1118.8	47.5	1865	12 ADI26144	Adi26144 Human CDN
14	1097.2	46.6	1446	12 ADO07812	Ado07812 Mouse pol
15	1097	46.6	21234	6 AAD32039	Aad32039 Human kin
16	813.2	34.5	1455	12 ADO07810	Ado07810 Human pol
17	811.6	34.5	1834	12 ADH58708	Adh58708 Human CDC
18	811.6	34.5	1834	13 ADRI4000	Adri4000 Human NF-
19	811.6	34.5	1834	13 ADP24718	Adp24718 PRO poly
20	768.8	32.7	1452	12 ADO07813	Ado07813 Mouse pol

21	768.4	32.6	906	12 ADO07814	Ado07814 Mouse pol
22	669.2	28.4	1785	10 ADG74696	Adg74696 Human kin
23	644.4	27.4	1743	12 ADH58715	Adh58715 Human CDC
24	643.2	27.3	1750	6 ABK84044	Abk84044 Human CDN
25	643.2	27.3	1750	10 ADL24750	Adl24750 Intestina
26	640.6	27.2	2254	12 ADE77073	Ade77073 Human CDN
27	640	27.2	2516	3 AAF21734	Aaf21734 Human bre
28	630.6	26.8	1643	10 ADL24751	Adl24751 Intestina
29	499.6	21.2	1500	10 ACAG3031	Acag3031 Human clk
30	499.6	21.2	1973	6 ABL64388	Abi64388 Stomach c
31	499.6	21.2	1973	6 ABN95692	Abn95692 Gene #219
32	499.6	21.2	1973	12 ADQ15173	Adq15173 Human can
33	499.6	21.2	1973	12 ADQ83230	Adq83230 Human tum
34	499.6	21.2	1973	12 ADQ83910	Adq83910 Human tum
35	499.6	21.2	1973	12 ADQ84393	Adq84393 Human tum
36	499.6	21.2	1973	12 ADQ85141	Adq85141 Human tum
37	499.6	21.2	1973	12 ADQ86323	Adq86323 Human tum
38	499.6	21.2	2111	9 AAD57339	Aad57339 Human kin
39	464.4	19.7	1296	4 AAH46906	Aah46906 CDNA enco
40	464.4	19.7	1842	5 AAS87541	Aas87541 DNA enco
41	461.6	19.6	1026	12 ADO07807	Ado07807 Human pol
42	458.6	19.5	1885	12 ADJ62819	Adj62819 Human CDN
43	452.2	19.2	1473	12 ADO07811	Ado07811 Mouse pol
44	443	18.8	443	3 AAC81771	Aac81771 Human bea
45	441.4	18.8	2905	8 ACC46217	Acc46217 Human dit

ALIGNMENTS

RESULT 1
AAD32038
ID AAD32038 standard; cDNA; 2354 BP.

AC AAD32038;

DT 18-JUN-2002 (first entry)

DE Human kinase protein CDNA.

KW Human; kinase protein; enzyme; cytotatic; osteopathic; gene expression;
KW colon-moderately differentiated adenocarcinoma; chromosome mapping; gene;
KW drug screening; therapeutic; gene therapy; tissue typing; chromosome 5;
KW bone osteosarcoma; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT 5'UTR 1..32

FT CDS /tag= a

FT CDS 33..1370

FT CDS /tag= b

FT 3'UTR /product= "Human kinase protein"

FT 1371..2354

FT /tag= c

FT WO200216567-A2.

PD 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US026389.

XX 24-AUG-2000; 2000US-0227470P.

PR 19-MAR-2001; 2001US-00810671.

XX (APPL-) APPLERA CORP.

PI Yan C, Ye J, Ketchum KA, Di Francesco V, Beasley EM;

DR WPI; 2002-269354/31.

DR P-FSDB; AAE20170.

XX New human kinase proteins and nucleic acids, useful in drug screening

assays, identifying modulators of kinase activity or treating disorders characterized by absence or unwanted expression of the protein.
 Claim 4; Fig 1; 81pp; English.

The invention relates to isolated human kinase proteins and nucleic acids. The nucleic acid and peptide sequences can be used as models for the development of human therapeutic targets, aid in the identification of therapeutic proteins and serve as targets for the development of human therapeutic agents that modulate kinase activity in cells and tissues that express the kinase. The nucleic acids are useful as probes or primers, in constructing recombinant vectors, for expressing antigenic portions of the proteins, chromosome mapping, drug screening, testing an individual for a genotype, and for gene therapy in patients containing cells that are aberrant in kinase gene expression. The proteins may be used in drug screening assays, in the identification of compounds that modulate, stimulate or inhibit kinase activity, in pharmacogenomic analysis, in treating disorders characterised by an absence or unwanted expression of the protein (bone osteosarcoma, or colon-moderately differentiated adenocarcinoma), and in generating antibodies specific for the peptides. Such antibodies can be used to detect the protein in situ, in vitro, or in cell lysate or supernatant, to isolate and purify the proteins from host cells, pharmacogenomic analysis, tissue typing, and in inhibiting protein function. The present sequence is human protein cDNA. Human kinase protein gene is located on chromosome 5

Sequence 2354 BP; 783 A; 370 C; 472 G; 729 T; 0 U; 0 Other;

Query Match 100.0%; Score 2354; DB 6; Length 2354;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GCCAGCTGGGTACTTTAAACACATGCTCCATGTGCATCCCTCTCGAAGCTTCGCACT	60
DB	1	GCCAGCTGGGTACTTTAAACACATGCTCCATGTGCATCCCTCTCGAAGCTTCGCACT	60
QY	61	CTGTTGAAGAGGACACTCATCCAGTCAATATTTAGAACGAGGTCTTTGAATGAGCGAG	120
DB	61	CTGTTGAAGAGGACACTCATCCAGTCAATATTTAGAACGAGGTCTTTGAATGAGCGAG	120
QY	121	ATTATCGGGACGGAGATACGTTGACGAATACAGGAATGACTACTGTGGAAGGATATGTC	180
DB	121	ATTATCGGGACGGAGATACGTTGACGAATACAGGAATGACTACTGTGGAAGGATATGTC	180
QY	181	CTAGACATTATCACAGAGACATTTGAAGCGGTATCGAATCCACTGCAAGTAATCTTCAG	240
DB	181	CTAGACATTATCACAGAGACATTTGAAGCGGTATCGAATCCACTGCAAGTAATCTTCAG	240
QY	241	TCCGAGCAGGAGACAGTCTCTAAAGGAGCGCAATAGACACTGTTCAAGTCATCAGT	300
DB	241	TCCGAGCAGGAGACAGTCTCTAAAGGAGCGCAATAGACACTGTTCAAGTCATCAGT	300
QY	301	CACGTTTCGAGAGCCACCGAAGAAAGATCCAGAGTATAGAGGATGATGAGGAGGTC	360
DB	301	CACGTTTCGAGAGCCACCGAAGAAAGATCCAGAGTATAGAGGATGATGAGGAGGTC	360
QY	361	ACCTGATCTGTCAGAGTGGAGCGTTCTTAAGAGCAAGATATGAATCGTGGACACTTTGG	420
DB	361	ACCTGATCTGTCAGAGTGGAGCGTTCTTAAGAGCAAGATATGAATCGTGGACACTTTGG	420
QY	421	GTGAGGAGCCCTTTGGCAAGTTGTAGAGTGCATTTGATCATGGCATGGATGGCATGCATG	480
DB	421	GTGAGGAGCCCTTTGGCAAGTTGTAGAGTGCATTTGATCATGGCATGGATGGCATGCATG	480
QY	481	TAGCAGTGAATCGTAAACATGTAGCGCGTTACCGTGAAGCAGCTGTTTCAGAAATCC	540
DB	481	TAGCAGTGAATCGTAAACATGTAGCGCGTTACCGTGAAGCAGCTGTTTCAGAAATCC	540
QY	541	AAGTATTAGACACTTAATAGTACTGATCCCAATAGTGTCTTCGATGTGTCCAGATGC	600
DB	541	AAGTATTAGACACTTAATAGTACTGATCCCAATAGTGTCTTCGATGTGTCCAGATGC	600
QY	601	TAGAATGTTTGTATCATGTGTCATGTTTGTATGTTGTAACACTACTGGGACTTAGTA	660

DB	601	TAGAATGTTTGTATCATGTGTCATGTTTGTATGTTTGAACACTACTGGGACTTAGTA	660
QY	661	CTTACGATTTTCAATTAAGAAACAGCTTTCTGCCATTTCAAAATGACACATCAGGAGA	720
DB	661	CTTACGATTTTCAATTAAGAAACAGCTTTCTGCCATTTCAAAATGACACATCAGGAGA	720
QY	721	TGGGATATCAGATCTGCGAGTCAATAAATTTTACATCATATAAATAAACCCATACAG	780
DB	721	TGGGATATCAGATCTGCGAGTCAATAAATTTTACATCATATAAATAAACCCATACAG	780
QY	781	ATCTGAAGCCCTGAAATATTTTGTGTGAAGTCTGACTATGATGATCAATATATTTCTA	840
DB	781	ATCTGAAGCCCTGAAATATTTTGTGTGAAGTCTGACTATGATGATCAATATATTTCTA	840
QY	841	AAATGAACCTGATGAAGCGACACTGAAACACAGATATCAAAAGTTGTTCAGCTTGGAA	900
DB	841	AAATGAACCTGATGAAGCGACACTGAAACACAGATATCAAAAGTTGTTCAGCTTGGAA	900
QY	901	GTGCAACGTATGATGATGAACATCAACAGTACTTTGGTGTCTACCCGGCAGCTACAGAG	960
DB	901	GTGCAACGTATGATGATGAACATCAACAGTACTTTGGTGTCTACCCGGCAGCTACAGAG	960
QY	961	CCGAGGTCATTTTGGCTTTAGGTTGGTCTCAGCCTTGTGATGTTTGGAGCATAGTTGCA	1020
DB	961	CCGAGGTCATTTTGGCTTTAGGTTGGTCTCAGCCTTGTGATGTTTGGAGCATAGTTGCA	1020
QY	1021	TTCTTATTTGAATATTTACCTTTTACAGTCTTTTACAGTCTCATGATAGTAAGAGCACC	1080
DB	1021	TTCTTATTTGAATATTTACCTTTTACAGTCTTTTACAGTCTCATGATAGTAAGAGCACC	1080
QY	1081	TGGCAATGATGAGAAAGCAATATTAGGACCCATACCAACACATGATTCAGAAAAAAGAA	1140
DB	1081	TGGCAATGATGAGAAAGCAATATTAGGACCCATACCAACACATGATTCAGAAAAAAGAA	1140
QY	1141	AACCAAGTATTTTACCATTAACAGCTAGATTGGGATGAACAGTCTCTGCTGGTAGAT	1200
DB	1141	AACCAAGTATTTTACCATTAACAGCTAGATTGGGATGAACAGTCTCTGCTGGTAGAT	1200
QY	1201	ATGTTAGGAGACGCTGCAACCGTTGAAGGAATTTATGCTTTGTCATGATCAAGAAACATG	1260
DB	1201	ATGTTAGGAGACGCTGCAACCGTTGAAGGAATTTATGCTTTGTCATGATCAAGAAACATG	1260
QY	1261	AGAACTGTTTTCAGCTGTTTCGAAAGATTTAGAAATATGATCCAACTCAAAAGATTAAC	1320
DB	1261	AGAACTGTTTTCAGCTGTTTCGAAAGATTTAGAAATATGATCCAACTCAAAAGATTAAC	1320
QY	1321	TGGATGAAGCATTTGAGCATCTTTTTCGAAAGATTTTGAATATGATCCAACTCAAAAG	1380
DB	1321	TGGATGAAGCATTTGAGCATCTTTTTCGAAAGATTTTGAATATGATCCAACTCAAAAG	1380
QY	1381	AGTGGTCTTACTATATCTCTCTAGAAAGATTTTAAAGTCTGTCAGTCAACTAAA	1440
DB	1381	AGTGGTCTTACTATATCTCTCTAGAAAGATTTTAAAGTCTGTCAGTCAACTAAA	1440
QY	1441	CATTCTAATATTTTGTAAACATTTAAATATTTTGTACAGTTAAGTGAATAATTTGTATG	1500
DB	1441	CATTCTAATATTTTGTAAACATTTAAATATTTTGTACAGTTAAGTGAATAATTTGTATG	1500
QY	1501	TTTTTGTATCAATAGCATTAATTAATCTTGTAGCAAGTATGCTTGTAAATGCAATGAA	1560
DB	1501	TTTTTGTATCAATAGCATTAATTAATCTTGTAGCAAGTATGCTTGTAAATGCAATGAA	1560
QY	1561	AAATTAATAATTTTCTTTTGAATTAACATTTTAAATACCTTTTGAATATCTCTT	1620
DB	1561	AAATTAATAATTTTCTTTTGAATTAACATTTTAAATACCTTTTGAATATCTCTT	1620
QY	1621	TGTGTCCAGTGAATAATGTGATCTTGTGCTTTGTACATGAGGTCACCTCTGAAGT	1680
DB	1621	TGTGTCCAGTGAATAATGTGATCTTGTGCTTTGTACATGAGGTCACCTCTGAAGT	1680
QY	1681	GATTTTCTTTCAGTAAAGGAAATCTTGACCTTATATTTTAAAGCAATATTTCTTTA	1740

Db 1681 GATTTTTCAGTAAAGGAAATCTTGACTACTTTATTTATCTTTAAAGGAATATCTTTTA 1740
Qy 1741 TATACCTCAAAATTTAGAACTTTAACTTTTAAAGTTTTCCTCTGTAATTTGTTGAACGGGTG 1800
Db 1741 TATACCTCAAAATTTAGAACTTTAACTTTTAAAGTTTTCCTCTGTAATTTGTTGAACGGGTG 1800
Qy 1801 ATTATTTAACTCTAGATAAGCAGGTACTAGAAACCAAACTCAGAAATGTTTACTGT 1860
Db 1801 ATTATTTAACTCTAGATAAGCAGGTACTAGAAACCAAACTCAGAAATGTTTACTGT 1860
Qy 1861 TAGAATCTTATTAATTTTAAAGTTTGTGTTTCTTTTTCATTTGGGTGATGTCAGGTGATA 1920
Db 1861 TAGAATCTTATTAATTTTAAAGTTTGTGTTTCTTTTTCATTTGGGTGATGTCAGGTGATA 1920
Qy 1921 ACCAGACATTCATGAAAGGATGAGTTTGTCCATTTGTGACAGTTTCTTTTAAATAAAC 1980
Db 1921 ACCAGACATTCATGAAAGGATGAGTTTGTCCATTTGTGACAGTTTCTTTTAAATAAAC 1980
Qy 1981 ACATACACATTTATTTAAAGATTTAAATCTAACTGGAAGTCAGTTTGGAAATGGAAT 2040
Db 1981 ACATACACATTTATTTAAAGATTTAAATCTAACTGGAAGTCAGTTTGGAAATGGAAT 2040
Qy 2041 TTCCAAGTATCTTGTGAGTCAAGATTAATAAATAGAAATTTCTGATGAGAGTTTCAG 2100
Db 2041 TTCCAAGTATCTTGTGAGTCAAGATTAATAAATAGAAATTTCTGATGAGAGTTTCAG 2100
Qy 2101 TTTTAAATACCAAGTCTTTAGGAGTCTTAAATCTGAGGAGATCTGTTTATCAAAATGACA 2160
Db 2101 TTTTAAATACCAAGTCTTTAGGAGTCTTAAATCTGAGGAGATCTGTTTATCAAAATGACA 2160
Qy 2161 TAAATACGTAACCTATTAAGATTTAAAGTTTATTAATTAAGGCAATTTATGTCGTGATAT 2220
Db 2161 TAAATACGTAACCTATTAAGATTTAAAGTTTATTAATTAAGGCAATTTATGTCGTGATAT 2220
Qy 2221 TCCTACGGGAGAAAGAGATTTGATTTGAAAGCAGTTTGGGAGAAAGTCTGCTGAAAT 2280
Db 2221 TCCTACGGGAGAAAGAGATTTGATTTGAAAGCAGTTTGGGAGAAAGTCTGCTGAAAT 2280
Qy 2281 TTCCAGAAATTTAAATGATTTGTTTACATAAACTTTTGTGACTTCAGAAAAAATAAAAA 2340
Db 2281 TTCCAGAAATTTAAATGATTTGTTTACATAAACTTTTGTGACTTCAGAAAAAATAAAAA 2340
Qy 2341 AACAAAAAATAAAC 2354
Db 2341 AACAAAAAATAAAC 2354

RESULT 2

AAF89402
ID AAF89402 standard; cDNA; 2497 BP.

AC AAF89402;

DT 14-AUG-2001 (first entry)

DE Human cell cycle regulating protein 53 coding sequence.

KW Human; cell cycle regulating protein 53; cancer; blood disease; HIV;
KW immunological disease; inflammation; ss.

OS Homo sapiens.

PN WO200130833-A1.

PD 03-MAY-2001.

PF 16-OCT-2000; 2000WO-CN000328.

PR 22-OCT-1999; 99CN-00119816.

PA (SHAN-) SHANGHAI BIO ROAD GENE DEV LTD.

PI Mao Y, Xie Y;

XX WPI; 2001-300480/31.
DR P-PSDB; RAB69705.

XX New cell cycle-regulating protein 53 and its polynucleotide, applicable
PT in diagnosis and treatment of malignant tumor; hemopathy, human
PT immunodeficiency virus infection, immunological diseases and various
PT inflammation.

XX Claim 6; Page 22-23; 30pp; Chinese.

XX The present invention provides the protein and coding sequences for human
CC cell cycle regulating protein 53. The sequences can be used in the
CC diagnosis and treatment of malignant tumours, haemopathy, human
CC immunodeficiency virus (HIV) infection, immunological diseases and
CC various types of inflammation. The present sequence is the cell cycle
CC regulating protein 53 coding sequence

XX Sequence 2497 BP; 801 A; 399 C; 542 G; 755 T; 0 U; 0 Other;

Query Match 94.1%; Score 2215.8; DB 4; Length 2497;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 2220; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 85 GTCATTTATTAAGAACGAGTCTTGAATGAGCGAGATTATCGGACCGGAGATACGTTG 144

Db 271 GTCATTTATTAAGAACGAGTCTTGAATGAGCGAGATTATCGGACCGGAGATACGTTG 330

Qy 145 ACGAATACAGGAATGACTACTGTCGAAGATATGTTCTTAGACATTTATCAGAGACATTTG 204

Db 331 ACGAATACAGGAATGACTACTGTCGAAGATATGTTCTTAGACATTTATCAGAGACATTTG 390

Qy 205 AAAGCGGTATCGAATCCACTGTCAGTAAATCTTTCAGTCCGAGCAGGAGAGACGTCTTA 264

Db 391 AAAGCGGTATCGAATCCACTGTCAGTAAATCTTTCAGTCCGAGCAGGAGAGACGTCTTA 450

Qy 265 AAAGAACGCAATAGACATCTGTTCAAGTCACTAGTCACTGTCGAAGACCCAGGAGCA 324

Db 451 AAAGAACGCAATAGACATCTGTTCAAGTCACTAGTCACTGTCGAAGACCCAGGAGCA 510

Qy 325 AAAGAACGCAATAGACATCTGTTCAAGTCACTAGTCACTGTCGAAGACCCAGGAGCA 384

Db 511 AAAGAACGCAATAGACATCTGTTCAAGTCACTAGTCACTGTCGAAGACCCAGGAGCA 570

Qy 385 TTCTAAGACGCAATAGACATCTGTTCAAGTCACTAGTCACTGTCGAAGACCCAGGAGCA 444

Db 571 TTCTAAGACGCAATAGACATCTGTTCAAGTCACTAGTCACTGTCGAAGACCCAGGAGCA 630

Qy 445 TAGAGTGCATTTGATCATGGCATGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCG 504

Db 631 TAGAGTGCATTTGATCATGGCATGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCG 690

Qy 505 TAGGCGGTACCGTGAAGCAGCTCTGTCAGAAATCCAGATTTAGAGCAGCTTAATAGTA 564

Db 691 TAGGCGGTACCGTGAAGCAGCTCTGTCAGAAATCCAGATTTAGAGCAGCTTAATAGTA 750

Qy 565 CTGATCCCAATAGTGTCTTCCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 624

Db 751 CTGATCCCAATAGTGTCTTCCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 810

Qy 625 ATGTTGTATTTGTTGAACTACTGGGACCTTAGTACTTAGTACTTAGTACTTAGTACTTAGT 684

Db 811 ATGTTGTATTTGTTGAACTACTGGGACCTTAGTACTTAGTACTTAGTACTTAGTACTTAGT 870

Qy 685 GCTTTCTGCCAATTTCAATTCAGACCATCAGGAGATGCGGTATCAGATCTGCCAGTCAA 744

Db 871 GCTTTCTGCCAATTTCAATTCAGACCATCAGGAGATGCGGTATCAGATCTGCCAGTCAA 930

Qy 745 TAAATTTTATACATCAATAATAATTAACCCATACAGATCTCGAAGCCTGAAATATTTTGT 804

Db 931 TAAATTTTATACATCAATAATAATTAACCCATACAGATCTCGAAGCCTGAAATATTTTGT 990

Qy 805 TTGTGAAGTCTGACTATGTGTAATTAATTTCTAAATGAACGTCATGAACGACAC 864

Db 991 TTGTGAAGTCTGACATATGTAGTCAAAATATAATTTCAAATGAAACGTGATGAACGCAC 1050
Qy 865 TGAATAACACAGATATCAAAAGTTGTGACTTTGGAGTGCACAGTATGATGATGACATC 924
Db 1051 TGAATAACACAGATATCAAAAGTTGTGACTTTGGAGTGCACAGTATGATGATGACATC 1110
Qy 925 ACAGTACTTTTGGTGTCTACCGGCACATACAGAGTCCCGAGTCAATTTTGGCTTTAGGTT 984
Db 1111 ACAGTACTTTTGGTGTCTACCGGCACATACAGAGTCCCGAGTCAATTTTGGCTTTAGGTT 1170
Qy 985 GGTCTCAGCCTTTGTGATGTTTGGAGCATAGGTTGCATTTCTTATGAAATATTACCTTGGTT 1044
Db 1171 GGTCTCAGCCTTTGTGATGTTTGGAGCATAGGTTGCATTTCTTATGAAATATTACCTTGGTT 1230
Qy 1045 TCACAGTCTTTTCAGACTCATGATAGTAAAGACCTCGCAATGATGGAACGAATATTAG 1104
Db 1231 TCACAGTCTTTTCAGACTCATGATAGTAAAGACCTCGCAATGATGGAACGAATATTAG 1290
Qy 1105 GACCCATACCAACACACATGATTTCAGAAAACAGAAAACGCAAGTATTTTTCCACCATAC 1164
Db 1291 GACCCATACCAACACACATGATTTCAGAAAACAGAAAACGCAAGTATTTTTCCACCATAC 1350
Qy 1165 AGCTAGATTTGGGATGAACACAGTCTCTGCTGTGTAGATATGTTAGGAGACGCTGCAACCGT 1224
Db 1351 AGCTAGATTTGGGATGAACACAGTCTCTGCTGTGTAGATATGTTAGGAGACGCTGCAACCGT 1410
Qy 1225 TGAAGGAATTTATGCTTTGTGATGATGAAGACATGAGAAATCTTTTGAACCTGTTGCA 1284
Db 1411 TGAAGGAATTTATGCTTTGTGATGATGAAGACATGAGAAATCTTTTGAACCTGTTGCA 1470
Qy 1285 GAATGTTAGATATGATCCAACTCAAGAAATTTACCTTCGATGAAGCATTTGCAGATCCCT 1344
Db 1471 GAATGTTAGATATGATCCAACTCAAGAAATTTACCTTCGATGAAGCATTTGCAGATCCCT 1530
Qy 1345 TCTTTGACTTTATTAAGAAAGAAATGAAATGGGAATCAGTGTCTTTACTATATCTTCTCT 1404
Db 1531 TCTTTGGCTTTATTAAGAAAGAAATGAAATGGGAATCAGTGTCTTTACTATATCTTCTCT 1590
Qy 1405 AGAAGAGATTACTTAAGACTGTGTCAGTCAACTAAGCATTTCTAATATTTTGTAAACAT 1464
Db 1591 AGAAGAGATTACTTAAGACTGTGTCAGTCAACTAAGCATTTCTAATATTTTGTAAACAT 1650
Qy 1465 AAATTAATTTGTACAGTTAAGTGAATATGATGTTTGTATCAATAGCATTAATTAAC 1524
Db 1651 AAATTAATTTGTACAGTTAAGTGAATATGATGTTTGTATCAATAGCATTAATTAAC 1710
Qy 1525 TTGTTAAGCAAGTATGCTTTGATATGATGCAATAGAAAAATTAATAATTTTCTTTT 1584
Db 1711 TTGTTAAGCAAGTATGCTTTGATATGATGCAATAGAAAAATTAATAATTTTCTTTT 1770
Qy 1585 GAAATTACATTTTAAATACCTTTGAAATATCTTTGTGTCCAGTGAATAATGTGATTG 1644
Db 1771 GAAATTACATTTTAAATACCTTTGAAATATCTTTGTGTCCAGTGAATAATGTGATTG 1830
Qy 1645 ATCTTGCTTTTGTACATGGAGGTCACTCTGAGTGAATTTTGTAGTAAAGGAAT 1704
Db 1831 ATCTTGCTTTTGTACATGGAGGTCACTCTGAGTGAATTTTGTAGTAAAGGAAT 1890
Qy 1705 CTTGACTACTTTTATATCTTTAAAGGAATTTCTTTATATATCTTCAAAATTTAGACTTAAC 1764
Db 1891 CTTGACTACTTTTATATCTTTAAAGGAATTTCTTTATATATCTTCAAAATTTAGACTTAAC 1950
Qy 1765 TTTTAAAGTTTTTCTTCTGTAATTTGTGAACGGGTGATTAATTAATCTAGATAAGCA 1824
Db 1951 TTTTAAAGTTTTTCTTCTGTAATTTGTGAACGGGTGATTAATTAATCTAGATAAGCA 2010
Qy 1825 GGTACTAGAAACCAAACTCAGAAAATGTTTACTGTTAGAAATTTCTATTAATTTTAAAGT 1884
Db 2011 GGTACTAGAAACCAAACTCAGAAAATGTTTACTGTTAGAAATTTCTATTAATTTTAAAGT 2070
Qy 1885 TTGATTTCTTTTTCATTCGGGTGATGTCAGGGTGAATAACGACATTCATGGAAGGCATG 1944

Db 2071 TTGATTTCTTTTTCATTCGGGTGATGTCAGGGTGAATAACGACATTCATGGAAGGCATG 2130
Qy 1945 CAGTTTGTCCCAATGTGACAGTTTCTTTAAATAAAACCATACACACTTTTATTAAGATTA 2004
Db 2131 CAGTTTGTCCCAATGTGACAGTTTCTTTAAATAAAACCATACACACTTTTATTAAGATTA 2190
Qy 2005 AAATCTTAACCTGGAAGTCAAGTCTTGGAATAATGCAATTTCCAAAGTATGTTTGGTGAAGTCA 2064
Db 2191 AAATCTTAACCTGGAAGTCAAGTCTTGGAATAATGCAATTTCCAAAGTATGTTTGGTGAAGTCA 2250
Qy 2065 AGATATAAAATAGAAATTTCTGATGAGAGGTTTCAAGTTTAAATACCAAGTCTTTAGAG 2124
Db 2251 AGATATAAAATAGAAATTTCTGATGAGAGGTTTCAAGTTTAAATACCAAGTCTTTAGAG 2310
Qy 2125 TCTTAAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAATACGTAACCTTATAGAAT 2184
Db 2311 TCTTAAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAATACGTAACCTTATAGAAT 2370
Qy 2185 AAGTTTATTAATTAGGCAATTTATGTCGTGATAATTTCTTACGGGAGAAAGAGGATTTGA 2244
Db 2371 AAGTTTATTAATTAGGCAATTTATGTCGTGATAATTTCTTACGGGAGAAAGAGGATTTGA 2430
Qy 2245 TTGAAAGCAGTTTGGGAAGAAAGTGTCTGCTGAAATTTCCAGAAATTAATTAATTTGTTGTTA 2304
Db 2431 TTGAAAGCAGTTTGGGAAGAAAGTGTCTGCTGAAATTTCCAGAAATTAATTAATTTGTTGTTA 2490
Qy 2305 CATAAAC 2311
Db 2491 CATAAAC 2497

RESULT 3
AAx89852
ID AAx89852 standard; cdna; 2446 BP.
XX
AC AAx89852;
XX
DT 08-OCT-1999 (first entry)
XX
DE Human protein kinase (HPKM)-3 encoding nucleotide (clone ID 339963).
XX
KW Human protein kinase molecule; HPKM; human; protein kinase;
phosphate group; cancer; immune disorder; ss.
XX
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 128..1483
FT /*tag= a
FT /product= "HPKM-3"
XX
PN WO9938981-A2.
XX
PD 05-AUG-1999.
XX
PF 12-JAN-1999; 99WO-US000661.
XX
PR 30-JAN-1998; 98US-00016000.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Bandman O, Hillman JL, Lal P, Akerblom IE, Shah P, Corley NC;
PI Guegler KJ;
XX
XX WPI; 1999-479190/40.
DR P-ESDB; AAY27054.
XX
PT New human protein kinase molecules useful for treating or preventing
cancer or an immune disorder.
XX
PS Claim 8; Page 74-75; 77pp; English.
XX
CC The invention provides human protein kinase molecules (HPKM) (AAY27052-

CC	57)	and nucleic acid sequences (AA989850-55) encoding the HPMK
CC	polypeptides respectively. The HPMK polypeptides can be produced	
CC	recombinantly by standard recombinant methodology. Protein kinases add	
CC	phosphate groups to proteins. HPMK polypeptide or an HPMK antagonist is	
CC	used to treat or prevent cancer or an immune disorder. The present	
CC	sequence represents the nucleotide sequence of HPMK-3	
XX		
SQ	Sequence 2446 BP; 788 A; 391 C; 516 G; 751 T; 0 U; 0 Other;	
	Query Match 87.2%; Score 2052.6; DB 2; Length 2446;	
	Best Local Similarity 95.8%; Pred. No. 0;	
	Matches 2155; Conservative 0; Mismatches 4; Indels 90; Gaps 1;	
Qy	85	GTCAATATTAGAAAGCAAGTCTTGTGAATGAGCGAGATATTACGGGACCGGAGATACGTTG 144
Db	288	GTCAATATTAGAAAGCAAGTCTTGTGAATGAGCGAGATATTACGGGACCGGAGATACGTTG 347
Qy	145	ACGAATACAGGAATGACTACTGTCAAGGATATGTTCTTAGACATTTATCACAGACACATTG 204
Db	348	ACGAATACAGGAATGACTACTGTGAAGATATGTTCTTAGACATTTATCACAGACACATTG 407
Qy	205	AAAGCGGGTATCGAATCCATCGCAGTAAATCTTCAGTCCGAGGAGGAGCAGTCCCTA 264
Db	408	AAAGCGGGTATCGAATCCATCGCAGTAAATCTTCAGTCCGAGGAGGAGCAGTCCCTA 467
Qy	265	AAAGGAAGCGCAATAGACACTGTTCAAGTCATCAGTCACGTTCCGAAGACCCGAAGGA 324
Db	468	AAAGGAAGCGCAATAGACACTGTTCAAGTCATCAGTCACGTTCCGAAGACCCGAAGGA 515
Qy	325	AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAGTGGAGACG 384
Db	516	-----AATCCGTGGACACTTTTGGGTGAAGGAGCCTTTGGCAAAAGTTG 557
Qy	385	TTCTAAGAGCAAGATATGAATCGTGGAACACTTTTGGGTGAAGGAGCCTTTGGCAAAAGTTG 444
Db	516	-----AATCCGTGGACACTTTTGGGTGAAGGAGCCTTTGGCAAAAGTTG 557
Qy	445	TAGAGTCATTTGATCGGCATGAGATGGCATGCATGTAGCGTGAAATCGTAAATAATG 504
Db	558	TAGAGTCATTTGATCATGAGCATGGCATGGCATGTAGCAGTGAAATCGTAAATAATG 617
Qy	505	TAGGCCGTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGACACCTTTAATAAGTA 564
Db	618	TAGGCCGTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGACACCTTTAATAAGTA 677
Qy	565	CTGATCCCAATAGTGTCTTCGATGTGTCCAGATGCTAGAAATGGTTTGATCATCATGGTC 624
Db	678	CTGATCCCAATAGTGTCTTCGATGTGTCCAGATGCTAGAAATGGTTTGATCATCATGGTC 737
Qy	625	ATGTTTGTATGTGTTTGAACTACGCGGACTTAGTACTTACGATTTTCATTAAGAAACA 684
Db	738	ATGTTTGTATGTGTTTGAACTACGCGGACTTAGTACTTACGATTTTCATTAAGAAACA 797
Qy	685	GCCTTCTGCCATTTCAAATGACCAATCAGGCGAGATGGCGTATCAGATCTGCCAGTCAA 744
Db	798	GCCTTCTGCCATTTCAAATGACCAATCAGGCGAGATGGCGTATCAGATCTGCCAGTCAA 857
Qy	745	TAAATTTTTTACATCAATAAATAAATTAACCCNATACAGATCTGAAGCCCTGAAAATATTTGT 804
Db	858	TAAATTTTTTACATCAATAAATAAATTAACCCNATACAGATCTGAAGCCCTGAAAATATTTGT 917
Qy	805	TTGTGAAGTCTGACTATGTAGTCAAAATATATTTCTAAATCAAACGATGACACGCACAC 864
Db	918	TTGTGAAGTCTGACTATGTAGTCAAAATATATTTCTAAATCAAACGATGACACGCACAC 977
Qy	865	TGAAAAACACAGATATCAAAAGTTGTTGACTTTTGAAGTGCAAACGATCATGATCAACATC 924
Db	978	TGAAAAACACAGATATCAAAAGTTGTTGACTTTTGAAGTGCAAACGATCATGATCAACATC 1037
Qy	925	ACAGTACTTTGGTGTCTACCCGCCACTACAGAGTCCCGAGGTCATTTTGGCTTTAGGTT 984
Db	1038	ACAGTACTTTGGTGTCTACCCGCCACTACAGAGTCCCGAGGTCATTTTGGCTTTAGGTT 1097

Db 1148 GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCAATTCCTTATTTGAATATTAACCTTGTTGTT 1207
Qy 1045 TCACAGTCTTTTCAGACTCATGATAGTAAGAGCACCTGGCAATGATGGAACGAATATTAG 1104
Db 1208 TCACAGTCTTTTCAGACTCATGATAGTAAGAGCACCTGGCAATGATGGAACGAATATTAG 1267
Qy 1105 GACCCATACCAACACATGATTCGAAACACGAAACCGCAAGTATTTTCCACCATACCC 1164
Db 1268 GACCCATACCAACACATGATTCGAAACACGAAACCGCAAGTATTTTCCACCATACCC 1327
Qy 1165 AGCTAGATTTGGGATGAAACACAGTCTCTGCTGTAGATATGTTAGGAGACGCTGCAAAACCGT 1224
Db 1328 AGCTAGATTTGGGATGAAACACAGTCTCTGCTGTAGATATGTTAGGAGACGCTGCAAAACCGT 1387
Qy 1225 TGAAGGAATTTATGCTTTGTGTCATGATGAAGAACATGAGAAACCTGTTTGACCTGGTTCGAA 1284
Db 1388 TGAAGGAATTTATGCTTTGTGTCATGATGAAGAACATGAGAAACCTGTTTGACCTGGTTCGAA 1447
Qy 1295 GAATGTTAGATATGATCCAACTCAAGAAATACCTTGGATGAAGCATTCGAGCATCCCTT 1344
Db 1448 GAATGTTAGATATGATCCAACTCAAGAAATACCTTGGATGAAGCATTCGAGCATCCCTT 1507
Qy 1345 TCTTTGACTTATTAAGAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATACCTTCTCT 1404
Db 1508 TCTTTGACTTATTAAGAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATACCTTCTCT 1567
Qy 1405 AGAAGAGATTACTTAAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATT 1464
Db 1568 AGAAGAGATTACTTAAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATT 1627
Qy 1465 AAATATTTTGTACAGTTAAGTGAATATTTGTATGTTTGTATCAATPAGCATATTAAC 1524
Db 1628 AAATATTTTGTACAGTTAAGTGAATATTTGTATGTTTGTATCAATPAGCATATTAAC 1687
Qy 1525 TTGTTAAGCAAGTAGTCTTGAATATGATTAATGCAATAGAAATTAATTAATTTTCTTTT 1584
Db 1688 TTGTTAAGCAAGTAGTCTTGAATATGATTAATGCAATAGAAATTAATTAATTTTCTTTT 1747
Qy 1585 GAAATATACATTTTAAATACCTTTGAATATTCCTTTGTGTCAGTGAATAATGTGATTG 1644
Db 1748 GAAATATACATTTTAAATACCTTTGAATATTCCTTTGTGTCAGTGAATAATGTGATTG 1807
Qy 1645 ATCTTGCTTTTGTACATGAGGTCACCTCTGAAGTGAATTTTGTGATGAATAAGGAAT 1704
Db 1808 ATCTTGCTTTTGTACATGAGGTCACCTCTGAAGTGAATTTTGTGATGAATAAGGAAT 1867
Qy 1705 CTGACTACTTTATATTTCTTAAAGAAATATCTTTATATATCTTCAATTTAGACTTAAC 1764
Db 1868 CTGACTACTTTATATTTCTTAAAGAAATATCTTTATATATCTTCAATTTAGACTTAAC 1927
Qy 1765 TTTAAAGTTTCTTCTGTAATTTGCAACGGGTGATTAATTAATTAATTAATTAATTAATTAAT 1824
Db 1928 TTTAAAGTTTCTTCTGTAATTTGCAACGGGTGATTAATTAATTAATTAATTAATTAATTAAT 1987
Qy 1825 GGTACTAGAAACCAAACTCAGAAATGTTTCTGTTAGAAATCTTATTAATTTTAAAGT 1884
Db 1988 GGTACTAGAAACCAAACTCAGAAATGTTTCTGTTAGAAATCTTATTAATTTTAAAGT 2047
Qy 1885 TTGATTTCTTTTCAATTCGGTGTGTCGGGTGATTAACGAGCATTCATGGAAGGCATG 1944
Db 2048 TTGATTTCTTTTCAATTCGGTGTGTCGGGTGATTAACGAGCATTCATGGAAGGCATG 2107
Qy 1945 CAGTTTGTCCATGTCAGATTTGTTTAAATAACCAATACACATCACTTTTATTAAGATTA 2004
Db 2108 CAGTTTGTCCATGTCAGATTTGTTTAAATAACCAATACACATCACTTTTATTAAGATTA 2167
Qy 2005 AAATCTAACTGGAAGTCTGAGTGGAAATGACATTTTCCAGTATGTTTGGTGAAGTCAAC 2064
Db 2168 AAATCTAACTGGAAGTCTGAGTGGAAATGACATTTTCCAGTATGTTTGGTGAAGTCAAC 2227
Qy 2065 AGATATAAAATAGAAATTCGATGAGAGGTTTCAGTTTAAATACCAAGTCCCTTAGAG 2124
Db 2228 AGATATAAAATAGAAATTCGATGAGAGGTTTCAGTTTAAATACCAAGTCCCTTAGAG 2287

Qy 2125 TCTTAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAAATACGTAAACCTTATAGAANT 2184
Db 2288 TCTTAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAAATACGTAAACCTTATAGAANT 2347
Qy 2185 AAGTTTATTAATTAGGCAATTTATGCTGTCGATATTTCTTACGGGAGAAAGAGGATTTGA 2244
Db 2348 AAGTTTATTAATTAGGCAATTTATGCTGTCGATATTTCTTACGGGAGAAAGAGGATTTGA 2407
Qy 2245 TTGAAAGACAGTTTGGGAGAAAGTGTGCTGCTGAAATTTCCAGAAATTTAATTTGATTGTTA 2304
Db 2408 TTGAAAGACAGTTTGGGAGAAAGTGTGCTGCTGAAATTTCCAGAAATTTAATTTGATTGTTA 2467
Qy 2305 CATAAACCTTTTGTACTTCA 2323
Db 2468 CATAAACCTTTTGTACTTCA 2486
RESULT 5
ADI29373
ID ADI29373 standard; cDNA; 2488 BP.
XX
AC ADI29373;
XX
DT 22-APR-2004 (first entry)
XX
Human MARK3-associated cDNA #40.
XX
Human; ss; antisense gene therapy; MARK3;
KW MAP/microtubule affinity-regulating kinase 3; cancer;
KW Alzheimer's disease; neurodegenerative disorder;
KW hyperproliferative disorder; cytostatic.
XX Homo sapiens.
XX US2003232771-A1.
XX
PD 18-DEC-2003.
XX
PF 17-JUN-2002; 2002US-00174319.
XX
PR 17-JUN-2002; 2002US-00174319.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Ward DT, Freier SM, Dobie KW;
XX
DR WPI; 2004-052188/05.
DR P-PSDB; ADI29255.
XX
New antisense compound targeted to a nucleic acid molecule encoding microtubule-affinity-regulating kinases (MARK3), useful for modulating expression of MARK3 or for treating cancer or Alzheimer's disease.
XX
PS Disclosure; Fig 2; 233pp; English.
XX
The invention relates to a compound comprising a sequence comprising 8-80 base pairs (bp) targeted to a nucleic acid encoding MARK3 (MAP/microtubule affinity-regulating kinase 3), that specifically hybridizes with the nucleic acid encoding MARK3 and inhibits expression of MARK3, i.e. is an antisense oligonucleotide (AO). Also included are a composition comprising the compound and a carrier or diluent, inhibiting the expression of MARK3 in cells or tissues, treating an animal having or suspected of having a disease or condition associated with MARK3 and screening for an antisense compound. The antisense oligonucleotide is useful for preparing a composition for treating hyperproliferative disorder, particularly cancer and neurodegenerative diseases e.g. Alzheimer's disease. The present sequence is a MARK3 associated cDNA included in the figures but not mentioned anywhere else in the specification.
XX
SQ Sequence 2488 BP; 819 A; 400 C; 504 G; 764 T; 0 U; 1 Other;

Query Match	87.0%	Score 2048;	DB 12;	Length 2488;
Best Local Similarity	95.9%	Pred. No. 0;		
Matches 2148;	Conservative 0;	Mismatches 1;	Indels 90;	Gaps 1;
QY	85	GTCAATTATTAGAACGAGTCCCTTGAATCAGCGAGATATCGGACCGGAGATACGTTG	144	
DB	338	GTCAATTATTAGAACGAGGATCCCTTGAATGAGCGAGATATCGGACCGGAGATACGTTG	397	
QY	145	ACGAATACAGGAATGACTACTGTGGAAGGATATGTTCTTAGACATATATCAGAGACATTTG	204	
DB	398	ACGAATACAGGAATGACTACTGTGGAAGGATATGTTCTTAGACATATATCAGAGACATTTG	457	
QY	205	AAAGCGGTATCGAATCCACTGCACTGAAATCTTTCACTCCGAGCAGGAGAACGATCCTTA	264	
DB	458	AAAGCGGTATCGAATCCACTGCACTGAAATCTTTCACTCCGAGCAGGAGAACGATCCTTA	517	
QY	265	AAAGGAGCGCATAGACACTGTTCAAGTCACTCAGTCCAGTCCGAGCAGGAGAACGATCCTTA	324	
DB	518	AAAGGAGCGCATAGACACTGTTCAAGTCACTCAGTCCAGTCCGAGCAGGAGAACGATCCTTA	561	
QY	325	AAAGATCCAGGAGTATAGAGATGATGAGGAGGTCACCTGATCTGTCAAAGTGGAGACG	384	
DB	562	-----	561	
QY	385	TTCTAAGACGAAGATGAATCGTGACACTTTGGGTGAAGAGCCTTTGGCAAGTTG	444	
DB	562	-----NATGAATCGTGACACTTTGGGTGAAGAGCCTTTGGCAAGTTG	607	
QY	445	TAGAGTGCAATGATCGGATGCGATGCGATGCACTGAGCAGTGAAATCGTAAAAAATG	504	
DB	608	TAGAGTGCAATGATCGGATGCGATGCGATGCACTGAGCAGTGAAATCGTAAAAAATG	667	
QY	505	TAGCCGTTACCGTGAAGCAGCTCGTTCAAGATCCAAAGTATTAGAGCCTTAAATAGTA	564	
DB	668	TAGCCGTTACCGTGAAGCAGCTCGTTCAAGATCCAAAGTATTAGAGCCTTAAATAGTA	727	
QY	565	CTGATCCCAATAGTGTCTCCGATGTGTCCAGATGCTAGAAATGGTTTGATCATCTGGTC	624	
DB	728	CTGATCCCAATAGTGTCTCCGATGTGTCCAGATGCTAGAAATGGTTTGATCATCTGGTC	787	
QY	625	ATGTTTGATGTGTTGAATCTAGTGGGACTTAGTACTTACGATTTCAATTAAGAAACA	684	
DB	788	ATGTTTGATGTGTTGAATCTAGTGGGACTTAGTACTTACGATTTCAATTAAGAAACA	847	
QY	685	GCCTTCGCCATTTCAAATTTGACACATCAGGAGATGGCGTATCAGATCTGCCAGTCAA	744	
DB	848	GCCTTCGCCATTTCAAATTTGACACATCAGGAGATGGCGTATCAGATCTGCCAGTCAA	907	
QY	745	TAAATTTTTTACATCAATAAATTAACCATACAGATCTGAAGCCTGAAATATTTTGT	804	
DB	908	TAAATTTTTTACATCAATAAATTAACCATACAGATCTGAAGCCTGAAATATTTTGT	967	
QY	805	TTGTGAAGTCTGACTATGTAGTCAATATAAATTTCTAAATTAAGAAACGATGAACACAC	864	
DB	968	TTGTGAAGTCTGACTATGTAGTCAATATAAATTTCTAAATTAAGAAACGATGAACACAC	1027	
QY	865	TGAAAAACACAGATATCAAAAGTTGTCACCTTTGGAAGTCAACGATGATGAACATC	924	
DB	1028	TGAAAAACACAGATATCAAAAGTTGTCACCTTTGGAAGTCAACGATGATGAACATC	1087	
QY	925	ACAGTACTTTTGGTCTTACCCGACATACAGAGCTCCCGAGGTCATTTTGGCTTTAGGTT	984	
DB	1088	ACAGTACTTTTGGTCTTACCCGACATACAGAGCTCCCGAGGTCATTTTGGCTTTAGGTT	1147	
QY	985	GGTCTCAGCCTTGTGATGTTTGAAGCATAGGTTGCACTTCTTATGAATATTAACCTTGGTT	1044	
DB	1148	GGTCTCAGCCTTGTGATGTTTGAAGCATAGGTTGCACTTCTTATGAATATTAACCTTGGTT	1207	
QY	1045	TCAAGTCTTTTCAAGTCAATGATGAAGACACCTGGCAATGATGAAGACGAATATTAG	1104	
DB	1208	TCAAGTCTTTTCAAGTCAATGATGAAGACACCTGGCAATGATGAAGACGAATATTAG	1267	
QY	1105	GACCCATPACCAACACATGATTCAGAAAAACAAGAAACGCAAGTATTTTTCACCAATAACC	1164	

Db 2348 AAGTTTATTATAGCAATTTATCTGTGTATTAATCTTACGGAGAAAGAGATTGA 2407
 Qy 2245 TTGGAACAGCTTTGGGAAGAAAGTCTGCTGAAATTTCCAGAAATTTAATTGATTGTTA 2304
 Db 2408 TTGGAACAGCTTTGGGAAGAAAGTCTGCTGAAATTTCCAGAAATTTAATTGATTGTTA 2467
 Qy 2305 CATAAACATTTTGACTTCA 2323
 Db 2468 CATAAACATTTTGACTTCA 2486
 RESULT 6
 ADP24393
 ID ADP24393 standard; cDNA; 1814 BP.
 AC ADP24393;
 XX
 DT 18-NOV-2004 (first entry)
 DE PRO polypeptide encoding cDNA SEQ ID NO:1571.
 XX ss; gene; PRO; antiinflammatory; antiarthritic; antirheumatic;
 KW immunosuppressive; osteopathic; antidiabetic; dermatological;
 KW antipsoriatic; antiallergic; antiasthmatic; hepatotropic; respiratory;
 KW gene therapy; immune system.
 XX Unidentified.
 XX WO2004041170-A2.
 PN 21-MAY-2004.
 XX 30-OCT-2003; 2003WO-US034312.
 PF 01-NOV-2002; 2002US-0423394P.
 PR (GETH) GENENTECH INC.
 PA Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
 PI Wu TD;
 XX WPI; 2004-419628/39.
 DR P-PSDB; ADP24394.
 XX
 PT New polypeptides and polynucleotides, useful for treating e.g.
 PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
 PT renal disease, or demyelinating diseases of the central or peripheral
 PT nervous system.
 PS Claim 1; SEQ ID NO 1571; 2940pp; English.
 XX
 CC The invention relates to a novel isolated nucleic acid and the PRO
 CC polypeptide encoded by it. A protein of the invention has
 CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
 CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
 CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
 CC of the invention may have a use in gene therapy. The PRO polypeptide, its
 CC agonist, antagonist, or antibody that specifically binds to the
 CC polypeptide is useful for treating an immune related disorder such as
 CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
 CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
 CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
 CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
 CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
 CC disease, a demyelinating disease of the central or peripheral nervous
 CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
 CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
 CC disease, infectious or autoimmune chronic active hepatitis, primary
 CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
 CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
 CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
 CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
 CC disease, asthma, allergic rhinitis, atopic dermatitis, food

CC hypersensitivity, urticaria, an immunologic disease of the lung,
 CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
 CC pneumonitis, a transplantation associated disease, graft rejection or
 CC graft-versus-host disease. The present sequence encodes a PRO protein of
 CC the invention.
 XX
 SQ Sequence 1814 BP; 609 A; 299 C; 385 G; 520 T; 0 U; 1 Other;
 Query Match 69.3%; Score 1632; DB 13; Length 1814;
 Best Local Similarity 99.3%; Pred. No. 5.8e-312;
 Matches 1638; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
 Qy 85 GTCATTATTAGAACCAAGTCTTGAATGAGCAGATATTCGGACCGGAGATACGTTG 144
 Db 161 GTCATTATTAGAACCAAGTCTTGAATGAGCAGATATTCGGACCGGAGATACGTTG 220
 Qy 145 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTATCACAGAGACATTG 204
 Db 221 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTATCACAGAGACATTG 280
 Qy 205 AAAGCGGTATCGAATCCACTGCAGTAAATCTTCAGTCCGAGCAGGAGAGACAGTCTTA 264
 Db 281 AAAGCGGTATCGAATCCACTGCAGTAAATCTTCAGTCCGAGCAGGAGAGACAGTCTTA 340
 Qy 265 AAAGGAAGCGCAATAGACACATGTTCAAGTCACTAGTCAGTTCGAGAGACCCAGGA 324
 Db 341 AAAGGAAGCGCAATAGACACATGTTCAAGTCACTAGTCAGTTCGAGAGACCCAGGA 400
 Qy 325 AAAGTCCAGGATATAGAGATGATGAGAGGGTCACTGTCTGCAAGTGTGAGAGACG 384
 Db 401 AAAGTCCAGGATATAGAGATGATGAGAGGGTCACTGTCTGCAAGTGTGAGAGACG 460
 Qy 385 TTCTAAGAGCAAGATATGAAATCGTGGACACTTTGGGTGAAGAGAGCCCTTGGCAAGTTG 444
 Db 461 TTCTAAGAGCAAGATATGAAATCGTGGACACTTTGGGTGAAGAGAGCCCTTGGCAAGTTG 520
 Qy 445 TAGAGTGCATTGATCATGGCATGGCATGTCATGTAGCAGTGAATTCGTAATAAATG 504
 Db 521 TAGAGTGCATTGATCATGGCATGGCATGTCATGTAGCAGTGAATTCGTAATAAATG 580
 Qy 505 TAGCCCGTTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTA 564
 Db 581 TAGCCCGTTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTA 640
 Qy 565 CTGATCCCAATAGTGTCTTCGATGTGTCCAGATGTCTAGATGTTGTTGATCATCATGTC 624
 Db 641 CTGATCCCAATAGTGTCTTCGATGTGTCCAGATGTCTAGATGTTGTTGATCATCATGTC 700
 Qy 625 ATGTTTGTATTGTTGTTGAACTACTGGGACTTTAGTACTTACGATTTCAATTAAGAAAACA 684
 Db 701 ATGTTTGTATTGTTGTTGAACTACTGGGACTTTAGTACTTACGATTTCAATTAAGAAAACA 760
 Qy 685 GCTTTCTGCCATTTCAATTCACACATCAGCAGATGGCGTATCAGATCTGCCAGTCAA 744
 Db 761 GCTTTCTGCCATTTCAATTCACACATCAGCAGATGGCGTATCAGATCTGCCAGTCAA 820
 Qy 745 TAAATTTTACATCAATAAATTAACCCATACAGATCTGAAAGCCTGAAATATTTTGT 804
 Db 821 TAAATTTTACATCAATAAATTAACCCATACAGATCTGAAAGCCTGAAATATTTTGT 880
 Qy 805 TTGTGAAGTCTGATATGATGTAATAAATTAATTCCTAATAAGAACGATGATGAACACAC 864
 Db 881 TTGTGAAGTCTGATATGATGTAATAAATTAATTCCTAATAAGAACGATGATGAACACAC 940
 Qy 865 TGAAAAACACAGATATCAAGTCTTGAGCTTTGGAAGTGAAGTCAAGTATCATGATGAACATC 924
 Db 941 TGAAAAACACAGATATCAAGTCTTGAGCTTTGGAAGTGAAGTCAAGTATCATGATGAACATC 1000
 Qy 925 ACAGTACTTTGGTGTCTACCCGGCAGTACAGAGTCCCGAGTCAATTTTGGCTTTAGGTT 984
 Db 1001 ACAGTACTTTGGTGTCTACCCGGCAGTACAGAGTCCCGAGTCAATTTTGGCTTTAGGTT 1060
 Qy 985 GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTGCAATTTTATTTGAATATTACCTTGGTT 1044

Db 1061 GGTCTCAGCCTTGTGATGTTTGGAGCATAGTGTGCATTTCTATTGAATATTACCTTGTT 1120
Qy 1045 TCACAGTCTTTTCAGACTCATGATAGTAAGACACCTGGCAATGATGGAACGATATTAG 1104
Db 1121 TCACAGTCTTTTCAGACTCATGATAGTAAGACACCTGGCAATGATGGAACGATATTAG 1180
Qy 1105 GACCCATACCAACACACATGATTCAGAAAAACAAGAAACGCAAGTATTTCACCATAC 1164
Db 1181 GACCCATACCAACACACATGATTCAGAAAAACAAGAAACGCAAGTATTTCACCATAC 1240
Qy 1165 AGCTAGATTGGGATGAACACAGTCTCTGCTGGTAGATATGTTAGGAGCGCTGCAAAACCGT 1224
Db 1241 AGCTAGATTGGGATGAACACAGTCTCTGCTGGTAGATATGTTAGGAGCGCTGCAAAACCGT 1300
Qy 1225 TGAAGGAATTTATGCTTGTGATGATGAAGAACATGAGAACTGTTGACCTGGTTGCA 1284
Db 1301 TGAAGGAATTTATGCTTGTGATGATGAAGAACATGAGAACTGTTGACCTGGTTGCA 1360
Qy 1285 GAATGTTAGATATGATCCAACTCAAGAAATTTACCTGGATGAAGCATTTGCAGCATCCTT 1344
Db 1361 GAATGTTAGATATGATCCAACTCAAGAAATTTACCTGGATGAAGCATTTGCAGCATCCTT 1420
Qy 1345 TCTTTGACTTATTAATAAGAAATGAATGGGAATCAGTGTCTTACTATATACTTCTCT 1404
Db 1421 TCTTTGACTTATTAATAAGAAATGAATGGGAATCAGTGTCTTACTATATACTTCTCT 1480
Qy 1405 AGAAGAGATTTACTTAAGACTGTGTCAGTCAACTAAACATTCATATATTTTGTAAACATT 1464
Db 1481 AGAAGAGATTTACTTAAGACTGTGTCAGTCAACTAAACATTCATATATTTTGTAAACATT 1540
Qy 1465 AAATTTATTTCTGACAGTTAAGTAAATATTGTATGTTTCTATCAATAGCATATTAAC 1524
Db 1541 AAATTTATTTCTGACAGTTAAGTAAATATTGTATGTTTCTATCAATAGCATATTAAC 1600
Qy 1525 TTGTTAAGCAAGTATGCTTGTGATAATGCAATAGAAAAATTTAAATTTTCTTTT 1584
Db 1601 TTGTTAAGCAAGTATGCTTGTGATAATGCAATAGAAAAATTTAAATTTTCTTTT 1660
Qy 1585 GAAATACCATTTTAAATACCTTTGAATATCCTTTGTGTCAGTGAATATGATG 1644
Db 1661 GAAATACCATTTTAAATACCTTTGAATATCCTTTGTGTCAGTGAATATGATG 1720
Qy 1645 ATCTTGCCCTTTGTACATGGAGGTCACTCTGAAGTGAATTTTGTAGTAAAGGAAT 1704
Db 1721 ATCTTGCCCTTTGTACATGGAGGTCACTCTGAAGTGAATTTTGTAGTAAAGGAAT 1780
Qy 1705 CTTGACTACTTTATATCTTTAAAGGAATA 1733
Db 1781 CTTGACTACTTTTAAAAAATAAAAAAAAAA 1809

RESULT 7

AD126142
ID AD126142 standard; cDNA; 1792 BP.

XX AC AD126142;

XX XX

DT 22-APR-2004 (first entry)

DE Human cDNA encoding protein that promotes STAT6 activation #54.

KW ss; gene; human; signal transducer and activator of transcription 6;
KW STAT6; immunogen; STAT6 activation; allergy; inflammation; cancer;
KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.

OS Homo sapiens.

XX WO2003104277-A2.

FN

XX

PD 18-DEC-2003.
XX 05-JUN-2003; 2003WO-JP007123.
PF 05-JUN-2002; 2002JP-00164257.
XX 06-JUN-2002; 2002US-0385912P.
PR 26-DEC-2002; 2002JP-043677326.
PR 27-DEC-2002; 2002US-0436467P.
PR 15-MAY-2003; 2003JP-00137505.
PR 16-MAY-2003; 2003US-0470836P.
XX (ASAH) ASahi KASEI KK.
PA Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;
PI WPI; 2004-122214/12.
XX P-PSDB; ADI26143.
DR New signal transducer and activator of transcription 6 activation
DR promoting purified protein, for diagnosing and treating disease
PT associated with activation/inhibition of transcription factor e.g.
PT diabetes and cancer.
PT
XX Claim 4; SEQ ID NO 107; 1368pp; English.
FS The invention relates to a purified protein promoting signal transducer
XX and activator of transcription 6 activation (STAT6). The protein is
CC useful for the producing an antibody, which involves administering the
CC protein or its epitope-bearing fragments to a non-human animal as an
CC antigen. The nucleic acid is useful for diagnosing a disease or
CC susceptibility to a disease related to expression or activity of the
CC protein. A transformant expressing the protein is useful for screening
CC compounds which inhibit or promote STAT6 activation. A transformant
CC expressing the protein is useful for producing a pharmaceutical
CC composition. Compositions, antibodies and antisense molecules are useful
CC for the treating a disease associated with STAT6 activation such as
CC allergic diseases, inflammation, autoimmune diseases, diabetes,
CC hyperlipidaemia, infections disease and cancers. Compositions are useful
CC for treating disease associated with STAT6 activation and/or prevention
CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
CC viral hepatitis and AIDS. The protein has efficient promoting STAT6
CC activity. The protein or nucleic acid is effectively useful for screening
CC compounds for treating and preventing disease associated with excessive
CC activation or inhibition of STAT6. The present sequence represents a
CC human cDNA encoding a protein which promotes STAT6 activation.
XX SQ Sequence 1792 BP; 587 A; 299 C; 385 G; 520 T; 0 U; 1 Other;

Query Match 69.2%; Score 1629.4; DB 12; Length 1792;

Best Local Similarity 99.9%; Pred. No. 1.9e-311;

Matches 1630; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 85 GTCATTATTTAGAAGCAAGTCTTGAATGACGAGATATTGGGACCGGAGATACGTTG 144

Db 161 GTCATTATTTAGAAGCAAGTCTTGAATGACGAGATATTGGGACCGGAGATACGTTG 220

Qy 145 ACGAATACAGGAATGACTACTGTGAAGATATGTTCTTAGACATTTATCACAGACATTTG 204

Db 221 ACGAATACAGGAATGACTACTGTGAAGATATGTTCTTAGACATTTATCACAGACATTTG 280

Qy 205 AAAGCGGTATCGAATCCACTGCCAGTAAATCTTCAGTCCGACGAGGAGAACGTCCTTA 264

Db 281 AAAGCGGTATCGAATCCACTGCCAGTAAATCTTCAGTCCGACGAGGAGAACGTCCTTA 340

Qy 265 AAAGGAAGCGCAATAGACATCTGTTCAAGTCATCATGTCAGTTCGAAGACCCGGAAGA 324

Db 341 AAAGGAAGCGCAATAGACATCTGTTCAAGTCATCATGTCAGTTCGAAGACCCGGAAGA 400

Qy 325 AAAGTCCAGGATATAGAGGATGATGAGGGGTCACTGATCTGTCAAAAGTGAGACG 384

Db 401 AAAGTCCAGGATATAGAGGATGATGAGGGGTCACTGATCTGTCAAAAGTGAGACG 460

Db 1. TTCAATTTGACCATCATCAGGCGAGATGGGTATCAGATCTGCCAGTCAATAAATTTTTTAC 60
Qy 757 ATCATATAAATAAATTAACCATACAGATCTGAAGCCCTGAAATAATTTTGTGTGAGTCTG 816
Db 61 ATCATATAAATAAATTAACCATACAGATCTGAAGCCCTGAAATAATTTTGTGTGAGTCTG 120
Qy 817 ACTATGTAGTCAATATAAATTTCTAAATGAACGCTGATGAACGACACTGAAAAACACAG 876
Db 121 ACTATGTAGTCAATATAAATTTCTAAATGAACGCTGATGAACGACACTGAAAAACACAG 180
Qy 877 ATATCAAAAGTTGTGACTTTTGGAAAGTGAACGCTGATGAACGACACTGAACTACTTTGG 936
Db 181 ATATCAAAAGTTGTGACTTTTGGAAAGTGAACGCTGATGAACGACACTGAACTACTTTGG 240
Qy 937 TGTCTACCCGGCACTACAGAGCTCCGAGGTCATTTTGGCTTTAGTTGGTCTCAGCCTT 996
Db 241 TGTCTACCCGGCACTACAGAGCTCCGAGGTCATTTTGGCTTTAGTTGGTCTCAGCCTT 300
Qy 997 GTGATGTTTGGAGCATAGTTTGCATTTCTTATTTGAATATTAATCTTGGTTTTCAGTCTTTC 1056
Db 301 GTGATGTTTGGAGCATAGTTTGCATTTCTTATTTGAATATTAATCTTGGTTTTCAGTCTTTC 360
Qy 1057 AGACTCATGATAGTAAAGAGCACCTGGCAATGATGAACGAAATATTAGGACCCATACAC 1116
Db 361 AGACTCATGATAGTAAAGAGCACCTGGCAATGATGAACGAAATATTAGGACCCATACAC 420
Qy 1117 AACACATGATTCAGAAAAACAGAAAAACGCAAGTATTTTCCACATAACGAGCTAGATTGGG 1176
Db 421 AACACATGATTCAGAAAAACAGAAAAACGCAAGTATTTTCCACATAACGAGCTAGATTGGG 480
Qy 1177 ATGAACACAGTTCTCTGCTGATAGTATGTTAGGAGCGCTGCAACCGTTGAAGAAATTTA 1236
Db 481 ATGAACACAGTTCTCTGCTGATAGTATGTTAGGAGCGCTGCAACCGTTGAAGAAATTTA 540
Qy 1237 TGCCTTTGTATGATGAAGAACATGAGAACTGTTTGAACCTGTTGCAAGAAATTTAGAT 1296
Db 541 TGCCTTTGTATGATGAAGAACATGAGAACTGTTTGAACCTGTTGCAAGAAATTTAGAT 600
Qy 1297 ATGATCCAACTCAAGAAATTAACCTTGGATGAGCAATGCAAGCATCTTCTTTGACTTAT 1356
Db 601 ATGATCCAACTCAAGAAATTAACCTTGGATGAGCAATGCAAGCATCTTCTTTGACTTAT 660
Qy 1357 TAAAAAGAAATGAATGGAAATCAGTGGTCTTACTATATCTCTCTAGAGAGATTAAC 1416
Db 661 TAAAAAGAAATGAATGGAAATCAGTGGTCTTACTATATCTCTCTAGAGAGATTAAC 720
Qy 1417 TTAAGACTGTGTCAGTCAACATAAATCTTAATATTTTGTAAACATTAATATTTTGT 1476
Db 721 TTAAGACTGTGTCAGTCAACATAAATCTTAATATTTTGTAAACATTAATATTTTGT 780
Qy 1477 ACAGTTAAGTGAATATTTGATGTTTGTATCAATAGCATTAATTAACCTGTTAAGCAAG 1536
Db 781 ACAGTTAAGTGAATATTTGATGTTTGTATCAATAGCATTAATTAACCTGTTAAGCAAG 840
Qy 1537 TATGCTCTTGTATAATGCAATAGAAAAATTAATTAATTTTCTTTTGAATATACCATT 1596
Db 841 TATGCTCTTGTATAATGCAATAGAAAAATTAATTAATTTTCTTTTGAATATACCATT 900
Qy 1597 TTTAAATACCTTTGAAATATCTTTGTCGCAAGTGAATATGATGATCTTGGCTTTT 1656
Db 901 TTTAAATACCTTTGAAATATCTTTGTCGCAAGTGAATATGATGATCTTGGCTTTT 960
Qy 1657 GTACATGGAGGTCACCTCTGAGTGTATTTTGTAGTAAAGAAATCTTGACTACTTT 1716
Db 961 GTACATGGAGGTCACCTCTGAGTGTATTTTGTAGTAAAGAAATCTTGACTACTTT 1020
Qy 1717 ATATCTTTAAAGGAATATTTCTTTATATACTTCAAAATTTAGAACTTAACTTTAAAGTTTT 1776
Db 1021 ATATCTTTAAAGGAATATTTCTTTATATACTTCAAAATTTAGAACTTAACTTTAAAGTTTT 1080
Qy 1777 TCTTCTGTAATTTGTTGAACGGGTGATATTTAATTAACCTTAGATAAGCAGTACTAGAAAC 1836
Db 1081 TCTTCTGTAATTTGTTGAACGGGTGATATTTAATTAACCTTAGATAAGCAGTACTAGAAAC 1140

Qy 1837 CAAAACTCAGAAAATGTTTACTGTAGAAATTTCTATTAAATTTTAAGTGTGTATCTTTT 1896
Db 1141 CAAAACTCAGAAAATGTTTACTGTAGAAATTTCTATTAAATTTTAAGTGTGTATCTTTT 1200
Qy 1897 TCATTGGGTGATGTCAGGGTGATAACAGACATTCATGGAAGGCATGCGAGTTTGTCCAT 1956
Db 1201 TCATTGGGTGATGTCAGGGTGATAACAGACATTCATGGAAGGCATGCGAGTTTGTCCAT 1260
Qy 1957 TGTGACAGTTTGTGTTTAAATAAACAACATACACACTTTTATTAAATTAAGATTAAATCTAACTGG 2016
Db 1261 TGTGACAGTTTGTGTTTAAATAAACAACATACACACTTTTAAATAAATAAATAAATAAATAA 1320
Qy 2017 AAA 2019
Db 1321 AAA 1323
RESULT 9
AD007809
ID AD007809 standard; cDNA; 1446 BP.
XX AD007809;
AC AD007809;
XX 01-JUL-2004 (first entry)
DT Human polynucleotide #39.
XX Human; gene; ss; fat cell number; fat cell size; obesity; diabetes;
KW anorectic; antidiabetic.
XX Homo sapiens.
OS US2004071700-A1.
PN 15-APR-2004.
XX 09-OCT-2002; 2002US-00267502.
PF 09-OCT-2002; 2002US-00267502.
PR (LIFE-) LIFE SCI DEV CORP.
PA Kim J, Galant R;
PI WPI; 2004-328526/30.
XX P-PSDB; ADO08026.
DR Identifying compounds that influence fat cell number or size for treating
XX or preventing obesity or diabetes by exposing the cell to the agent and
PT identifying fat cell number or size relative to cells not exposed to the
XX agent.
PS Claim 1; SEQ ID NO 135; 275pp; English.
XX The invention relates to a method of identifying compounds that influence
CC fat cell number or size comprising providing a cell that expresses a gene
CC and an agent, exposing the cell to the agent and identifying fat cell
CC number or size relative to cells not exposed to the agent. The method
CC also comprises providing an expression vector and an agent, exposing the
CC vector to the agent, detecting a change in expression of the gene
CC relative to expression of the gene in an expression vector not exposed to
CC the agent, treating a subject with the agent and identifying fat cell
CC number or size in the subject. The agent comprises a mammal, preferably a human. The
CC oligonucleotide. The subject comprises a polypeptide and an agent, exposing the
CC method also comprises providing a polypeptide of the agent to the
CC polypeptide to the agent, detecting binding of the agent to the
CC polypeptide or a change in an activity of the polypeptide, treating a
CC subject with the agent and identifying fat cell number or size in the
CC subject. The agent comprises an antibody. A method of regulating fat cell
CC number or size comprises providing a subject containing fat cells and an
CC agent that changes the expression of a gene, and treating the subject
CC with the agent under conditions so that fat cell size or number in the

CC subject is altered. The method is useful for identifying compounds that
 CC influence fat cell number or size, for preparing a composition for
 CC treating or preventing obesity or diabetes. This sequence represents
 CC human cDNA used in the scope of the invention.
 XX
 SQ Sequence 1446 BP; 474 A; 258 C; 332 G; 382 T; 0 U; 0 Other;

Query Match		54.6%;	Score 1286;	DB 12;	Length 1446;
Best Local Similarity		100.0%;	Pred. No. 8.8e-244;		
Matches 1286;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	85	GTCAATATTAGAACCAAGGCTCTTGAATGAGCGAGATATCGGACCGGAGATACGTTG	144		
DB	161	GTCAATATTAGAACCAAGGCTCTTGAATGAGCGAGATATCGGACCGGAGATACGTTG	220		
QY	145	ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTAGACATTATCACAGACATTG	204		
DB	221	ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTAGACATTATCACAGACATTG	280		
QY	205	AAAGCGGTATCGAATCCACTGACGTAATCTTCAGTCCGACGAGGAGACGAGCTCTTA	264		
DB	281	AAAGCGGTATCGAATCCACTGACGTAATCTTCAGTCCGACGAGGAGACGAGCTCTTA	340		
QY	265	AAAGGAAGCGCAATAGACACTGTTCAAGTCATCAGTCAGTTCGAAGACCCACCGAAGA	324		
DB	341	AAAGGAAGCGCAATAGACACTGTTCAAGTCATCAGTCAGTTCGAAGACCCACCGAAGA	400		
QY	325	AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAGTGAGACG	384		
DB	401	AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAGTGAGACG	460		
QY	385	TTCTAAGAGCAAGATATGAAATCGTGGACACTTTTGGGTGAAGGAGCCCTTGGCAAAGTTG	444		
DB	461	TTCTAAGAGCAAGATATGAAATCGTGGACACTTTTGGGTGAAGGAGCCCTTGGCAAAGTTG	520		
QY	445	TAGAGTGATTCATGATGGCATGGATGGATGCGATGCTAGTGTGAGAGTGAATAATCGTAA	504		
DB	521	TAGAGTGATTCATGATGGCATGGATGGATGCGATGCTAGTGTGAGAGTGAATAATCGTAA	580		
QY	505	TAGGCGGTATCCGTAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCACTTAATAGTA	564		
DB	581	TAGGCGGTATCCGTAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCACTTAATAGTA	640		
QY	565	CTGATCCCAATAGTGTCTTCCGATGTGTCAGATGCTAGAAATGGTTTGATCATCATGTC	624		
DB	641	CTGATCCCAATAGTGTCTTCCGATGTGTCAGATGCTAGAAATGGTTTGATCATCATGTC	700		
QY	625	ATGTTTGTATTTGTTTGAACCTACTGGGACTTAGTACTTACGATTTCAATTAAGAAACA	684		
DB	701	ATGTTTGTATTTGTTTGAACCTACTGGGACTTAGTACTTACGATTTCAATTAAGAAACA	760		
QY	685	GCTTCTGCGCATTTCAAATTTGACACATCAGGACATGCGGTATCAGATCTGCCAGTCAA	744		
DB	761	GCTTCTGCGCATTTCAAATTTGACACATCAGGACATGCGGTATCAGATCTGCCAGTCAA	820		
QY	745	TAAATTTTTTACATCATATAATTAATTAACCCATACAGATCTGAAGCCTGAAATATTTTGT	804		
DB	821	TAAATTTTTTACATCATATAATTAATTAACCCATACAGATCTGAAGCCTGAAATATTTTGT	880		
QY	805	TTGTGAAGTCTGATGATGTGATCAATATTAATTTCAAATTAAGAACGTCATGACGCAC	864		
DB	881	TTGTGAAGTCTGATGATGTGATCAATATTAATTTCAAATTAAGAACGTCATGACGCAC	940		
QY	865	TGAAAAACACAGATATCAAGTGTGTTGACTTTTGGAGTGCAACGTCATGATGACATC	924		
DB	941	TGAAAAACACAGATATCAAGTGTGTTGACTTTTGGAGTGCAACGTCATGATGACATC	1000		
QY	925	ACAGTACTTTGGTGTCTACCGGCACTACAGAGCTCCCGAGCTCAATTTTGGCTTTAGTT	984		
DB	1001	ACAGTACTTTGGTGTCTACCGGCACTACAGAGCTCCCGAGCTCAATTTTGGCTTTAGTT	1060		
QY	985	GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCAATCTTATTGAATATTACCTTGGTT	1044		

DB	1061	GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCAATCTTATTGAATATTACCTTGGTT	1120		
QY	1045	TCACAGTCTTTTCAGAGCTCATGATAGTAAAGAGCAGCTGGCAATGATGGAACGAATATTAG	1104		
DB	1121	TCACAGTCTTTTCAGAGCTCATGATAGTAAAGAGCAGCTGGCAATGATGGAACGAATATTAG	1180		
QY	1105	GACCCATACCAACACACATGATTTCAAGAAACGAAGAAACGCAAGTATTTCACCAATAACC	1164		
DB	1181	GACCCATACCAACACACATGATTTCAAGAAACGAAGAAACGCAAGTATTTCACCAATAACC	1240		
QY	1165	AGCTAGATTGGGATGAACACACAGTCTCTGCTGTAGATATGTTAGGAGAGCTGCAACCGT	1224		
DB	1241	AGCTAGATTGGGATGAACACACAGTCTCTGCTGTAGATATGTTAGGAGAGCTGCAACCGT	1300		
QY	1225	TGAAGGAATTTTATGCTTTTGTCTATGATGAAGAACATGAGAAACTGTTGACCTGGTTCGAA	1284		
DB	1301	TGAAGGAATTTTATGCTTTTGTCTATGATGAAGAACATGAGAAACTGTTGACCTGGTTCGAA	1360		
QY	1285	GAATGTTAGAATATGATCCAACTCAAAGAAATTAACCTTGGATGAAGCATTCAGAGCATCTT	1344		
DB	1361	GAATGTTAGAATATGATCCAACTCAAAGAAATTAACCTTGGATGAAGCATTCAGAGCATCTT	1420		
QY	1345	TCCTTGACTTATTAAAAAGGAATGA	1370		
DB	1421	TCCTTGACTTATTAAAAAGGAATGA	1446		
RESULT 10					
ID	AD126140	standard; cDNA; 1881 BP.			
AC	AD126140;				
DT	22-APR-2004	(first entry)			
XX	Human cDNA encoding protein that promotes STAT6 activation #53.				
DE	ss; gene; human; signal transducer and activator of transcription 6;				
KW	STAT6; immunogen; STAT6 activation; allergy; inflammation;				
KW	autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;				
KW	Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;				
KW	systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;				
KW	ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.				
OS	Homo sapiens.				
XX	WO2003104277-A2.				
PN	18-DEC-2003.				
XX	05-JUN-2003; 2003WO-JP007123.				
XX	05-JUN-2002; 2002JP-00164257.				
PR	06-JUN-2002; 2002US-0385912P.				
PR	26-DEC-2002; 2002JP-0037326.				
PR	27-DEC-2002; 2002US-0436467P.				
PR	15-MAY-2003; 2003JP-00137505.				
PR	16-MAY-2003; 2003US-0470836P.				
XX	(ASAH) ASahi KASEI KK.				
PA	Sugahara T, Matsuda A, Honda G, Muramatsu S, Iehizawa K;				
XX	WPI, 2004-122214/12.				
DR	P-FSDB; AD126141.				
XX	New signal transducer and activator of transcription 6 activation				
PT	promoting purified protein, for diagnosing and treating disease				
PT	associated with activation/inhibition of transcription factor e.g.				
PT	diabetes and cancer.				
XX	Claim 4; SEQ ID NO 105; 1368pp; English.				
PS					
XX					

The invention relates to a purified protein promoting signal transducer and activator of transcription 6 activation (STAT6). The protein is useful for producing an antibody, which involves administering the protein or its epitope-bearing fragments to a non-human animal as an antigen. The nucleic acid is useful for diagnosing a disease or susceptibility to a disease related to expression or activity of the protein. A transformant expressing the protein is useful for screening compounds which inhibit or promote STAT6 activation. A transformant expressing the protein is useful for producing a pharmaceutical composition. Compositions, antibodies and antisense molecules are useful for the treating a disease associated with STAT6 activation such as allergic diseases, inflammation, autoimmune diseases, diabetes, hyperlipidaemia, infectious disease and cancer. Compositions are useful for treating disease associated with STAT6 activation and/or prevention of Th1 hyperactive diseases. Compositions are also useful in rheumatoid arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma, allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage, viral hepatitis and AIDS. The protein has efficient promoting STAT6 activity. The protein or nucleic acid is effectively useful for screening compounds for treating and preventing disease associated with excessive activation or inhibition of STAT6. The present sequence represents a human cDNA encoding a protein which promotes STAT6 activation.

SQ Sequence 1881 BP; 570 A; 344 C; 450 G; 517 T; 0 U; 0 Other;
Query Match 54.3%; Score 1277.4; DB 12; Length 1881;
Best Local Similarity 87.9%; Pred No. 4.6e-242;
Matches 1413; Conservative 0; Mismatches 193; Indels 3; Gaps 2;

Qy 85 GTCATTATTAGAACGAAGGCTCTTGAATGACGAGATTATCGGACCGGAGATACGTTG 144
Db 250 GTCATTATTAGAACGAAGTGTGTAATGAGAGAGATTATCGGACCGGAGATACATTG 309
Qy 145 ACGAATACAGGAATGACTCTGTGAGGATATGTTCTTAGACATTATCACAGACATTG 204
Db 310 ATGAATACAGAAATGACTACTCGGAAGGATATGTTCCAAAGACATTACCATAGACGTTG 369
Qy 205 AAAGCGGTATCGAATCCACATCGAGTAATCTTCAGTCCGACGAGGAGACGATCCCTA 264
Db 370 AAAGCACTTACCGGATCCATTCGAGTAATCTTCAGTCCGACGAGGAGACGATCCCTA 429
Qy 265 AAAGGAACGGAATAGACATCTGTTCAAGTCATCATGTCACGTTGGAAGACCAACGGA 324
Db 430 AGAGAAGCGTAATAGACCTGTGCAAGTCATCATGTCGATTCGAAAGACCAACGGA 489
Qy 325 AAAGATCCAGAGTATAGAGTATGATGAGGAGTCACTGATCTGTCGAAAGTGAGAG 384
Db 490 AAAGATCCAGAGTATAGAGTATGATGAGGAGTCACTGATCTGTCGAAAGTGAGAG 549
Qy 385 TTCTAAGACGAAGATATGAATTCGGACATCTTGGGTGAAGGAGCCTTTGGCAAAAGTTG 444
Db 550 TTCTAAGACGAAGATATGAATTCGGACATCTTGGGTGAAGGAGCCTTTGGCAAAAGTTG 609
Qy 445 TAGAGTGCATTGATCATCGGCATGGATGCGATGTCATGATGAGCAATCGTAAAGAAATG 504
Db 610 TAGAGTGCATTGATCATCGGCATGGATGCGTACATGATGAGCAATCGTAAAGAAATG 669
Qy 505 TAGGCGGTACCGTCAAGCAGCTCGTTGAGAAATCCAGATATTAGACACTTAATAGTA 564
Db 670 TAGGCGGTACCGGAGGAGCAGCTCGTTGAGAAATCCAGATATTGAGCACTTGAACGCA 729
Qy 565 CTGATCCCAATAGTGTCTTCGATGTCGATGTCGATGATGATGATGATGATGATGATG 624
Db 730 CTGATCCCAATAGTGTCTTCGATGTCGATGTCGATGATGATGATGATGATGATGATG 789
Qy 625 ATGTTGTATTGTTGAACTA CTGGGACTTATGATGATGATGATGATGATGATGATGATG 684
Db 790 ATGTTGTATTGTTGAACTA CTGGGACTTATGATGATGATGATGATGATGATGATGATG 849
Qy 685 GCTTCTGCCATTTCAAAATTCACATCAGCAGATCGCGATGATGATGATGATGATGATG 744
Db 850 GTTTTCTGCCATTTCAAAATTCACATCAGCAGATGATGATGATGATGATGATGATGATG 909

Qy 745 TAAATTTTTACATCATATAAATAAATTAACCCATACAGATCTGAAGCCTGAATAATATTTTGT 804
Db 910 TAAATTTTTACATCATATAAATAAATAAATTAACACACACGACCTAAAACCTGAAATAATTTTAT 969
Qy 805 TTGTGAAGTCTGACTATGTAGTCAAAATATTAATCTTAATTAAGAAAGTGATGAACGACAC 864
Db 970 TTGTGAAGTCTGACTATGTAGTCAAAATATTAATCTTAATTAAGAAAGTGATGAACGACAC 1029
Qy 865 TGAATAACACAGATATCAAAAGTCTGTGACTTTTGAAGTGCAACGATGATGATGATGAACATC 924
Db 1030 TGAATAACACAGATATCAAAAGTCTGTGACTTTTGAAGTGCAACATGATGATGATGAACATC 1089
Qy 925 ACAGTACTTTGGTGTCTACCCGGACATACAGAGCTCCCGAGGTCAATTTTGGCTTTAGGTT 984
Db 1090 ATAGTACTTTGGTGTCTCCACAGGCACTACAGGGCTCCAGAGGTCAATTTTGGCTTTAGGTT 1149
Qy 985 GGTCTACGCTTGTGATGTTTGGAGCATAGTTTGCATCTTATTGAAATATTACCTTGGTT 1044
Db 1150 GGTCTACGCTTGTGATGTTTGGAGCATAGGCTGCACTTTATTGAGTACTACCTTGGTT 1209
Qy 1045 TCACAGTCTTTGACACTCATGATAGTAAAGAGCACCTGGCAATGATGAAACGAAATATTAG 1104
Db 1210 TCACAGTCTTTGACAGCCACGATAGTAAAGAGCACCTGGCAATGATGAGGCGGATCTTAG 1269
Qy 1105 GACCCATACCAACACATGATTTCAGAAAAACAAGAAACGCAAGTATTTTCCACATAACC 1164
Db 1270 GACCCATCCAGCACATATGATTCAGAAAGACAGAAACGCAAGTATTTTCCACATAACC 1329
Qy 1165 AGCTAGATTGGGATGAACACAGTTCTGCTGTAGATATGTTAGGAGCGCTGCAACCGT 1224
Db 1330 AGCTAGATTGGGATGAACATAGTTTCACTGGGAGATATGTTAGGAGCGCTGCAACCGT 1389
Qy 1225 TGAAGGAATTTATGCTTTGTCATGATGAAGAAACATGAAAACTGTTTCACTGCTGTCGAA 1284
Db 1390 TAAAGGAATTTATGCTGTGTCATGACGAGAGCATGAGAACTGTTTCACTGCTGTCGAA 1449
Qy 1285 GAATTTAGAAATATGATCCAACTCAAGAAATTA CTTGGATGAAGCATTTGAGCATTCCTT 1344
Db 1450 GAATTTGGAGTATGATCCAGCGAGAGGATCA CTTGGATGAAGCATTTGAGCATTCCTT 1509
Qy 1345 TCTTTGACTTATTAAGAAAGAAATGAATGGGAATCA GGTGCTTACTATATATCTTCTCT 1404
Db 1510 TCTTTGACTTATTAAGAAAGAAATGAAGGAAATG -AGTGGGAGTCAAGGCTTCTCTGTGTA 1568
Qy 1405 AGAAGAGTAATCTTAAGACTGTGTCACTCAACTAAACATTTCTAATATTTTGTAAACATT 1464
Db 1569 GGAGCAGTTACTTCCAGACTGTGTCACTCAACTAAACCTTCTAATATTTTGTAAACATT 1628
Qy 1465 AAATTTATTTGTA CAGTTAAGTAAATATTTGATGTTTGTATGATCAATAGCATTAATTAAC 1524
Db 1629 AAATTTATTTGTA CAGTTAAGTAAATATTTGATGTTTGTATGATCAATAGCATTAATTAAC 1688
Qy 1525 TTGTTAAGCAATGATGTTTGTATGATTAATGCAATGCAATTAAGAAATTAATTTTCTTTT 1583
Db 1689 TTGTTAAGTATGTTTGTATGATTAATGCAATGCAATTAAGAAATTAATTTTCTTTT 1748
Qy 1584 -TGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTTCTTTT 1642
Db 1749 ATGTTAATGCACTTTTAAGGCTTTTAGATGCTTTGTTGTTGTTGTTGTTGTTGTTGTTG 1808
Qy 1643 TGATCTTGGCTTTTGTACATGAGGTCACCTCTGAAGTGAATTTTTTTT 1690
Db 1809 TGGTCCCATCTTTTGTACATGAAGGTTGACTCTGAAGTGAATTTTTTTT 1856

RESULT 11
AAC59283
ID AAC59283 standard; cDNA; 1222 BP.
XX
AC AAC59283;
XX 02-FEB-2001 (first entry)
DT
XX

DE Human secreted protein cDNA #7.
 XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anti-ulcer;
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200056753-A1.
 XX
 XX 28-SEP-2000.
 XX
 XX 16-MAR-2000; 2000WO-US006765.
 XX
 XX 23-MAR-1999; 99US-0126051P.
 XX 10-DEC-1999; 99US-0169906P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Ruben SM, Komatsoulis G;
 XX
 XX WPI; 2000-594570/56.
 XX P-ESDB; AAB33724.
 XX
 XX Nucleic acid molecules encoding human secreted proteins, used in
 XX preventing, treating or ameliorating a disorder.
 XX
 XX Claim 1; Page 349-350; 410pp; English.
 XX
 XX The invention relates to the isolation of genes AAC59277-C59325 encoding
 XX 49 human secreted proteins AAB33718-B33764. The genes can be used to
 XX generate fusion proteins by linking to the gene for the human
 XX immunoglobulin G Fc portion (SEQID1) for increasing the stability of the
 XX fusion protein as compared to the human protein only. The genes and
 XX proteins are useful for preventing, ameliorating or treating medical
 XX conditions, e.g. by protein or gene therapy. The genes are isolated from
 XX a range of human tissues disclosed in the specification. The nucleic
 XX acids, proteins, antibodies and (ant)agonists are useful in the
 XX diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 XX ovarian cancer, and other cancers of the adrenal gland, bone, bone
 XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b)
 XX immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 XX cardiovascular disorders such as myocardial ischaemia; (d) wound healing
 XX ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 XX infectious diseases such as viral, bacterial, fungal and parasitic
 XX infections
 XX
 XX Sequence 1222 BP; 407 A; 194 C; 229 G; 392 T; 0 U; 0 Other;
 XX
 XX Query Match 50.3%; Score 1185; DB 3; Length 1222;
 XX Best Local Similarity 99.9%; Pred. No. 6.9e-224;
 XX Matches 1196; Conservative 0; Mismatches 1; Gaps 1;
 XX
 XX 517 GTGAAGCAGCTCGTTTCCAGAAATCCAAATATTAGAGCAGCTTAAATAGTACTATCCCAATA 576
 XX 9 GTGAAGCAGCTCGTTTCCAGAAATCCAAATATTAGAGCAGCTTAAATAGTACTATCCCAATA 68
 XX
 XX 577 GTGTCTCCCGATGTGTCAGAGATGCTAGAAATGGTTTGATCATCTGTCATGTTGTTGATTG 636
 XX 69 GTGTCTCCCGATGTGTCAGAGATGCTAGAAATGGTTTGATCATCTGTCATGTTGTTGATTG 128
 XX
 XX 637 TGTGTTGAATCTAGGAGCTTACTTACGATTTCAATTAAGAAACAGCTTTCTGCCAT 696
 XX
 XX 129 TGTGTTGAATCTAGGAGCTTACTTACGATTTCAATTAAGAAACAGCTTTCTGCCAT 188
 XX
 XX 697 TTTCAATTTGACCATCAGGAGATGGGATCAGATCTGCAGCTCAATTAATTTTTTAC 756
 XX
 XX 189 TTTCAATTTGACCATCAGGAGATGGGATCAGATCTGCAGCTCAATTAATTTTTTAC 248

QY 757 ATCATATAAATTAACCCATACAGATCTGAAGCCTGAAAAATATTTTGTGTAAGTCTG 816
 DB 249 ATCATATAAATTAACCCATACAGATCTGAAGCCTGAAAAATATTTTGTGTAAGTCTG 308
 QY 817 ACTATGATGCAATATAAATTTCTAAATGAAACGATGAACGACACTGAAAAACACAG 876
 DB 309 ACTATGATGCAATATAAATTTCTAAATGAAACGATGAACGACACTGAAAAACACAG 368
 QY 877 ATATCAAAAGTTGTTGACTTTTGGAAAGTCAACGATGATGATGAACATCACAGTACTTTGG 936
 DB 369 ATATCAAAAGTTGTTGACTTTTGGAAAGTCAACGATGATGATGAACATCACAGTACTTTGG 428
 QY 937 TGTCTACCCGCACTACAGAGCTCCGAGGTCATTTTGGCTTTAGGTTGGTCTCAGCCTT 996
 DB 429 TGTCTACCCGCACTACAGAGCTCCGAGGTCATTTTGGCTTTAGGTTGGTCTCAGCCTT 488
 QY 997 GTGATGTTGGAGCATAGGTTGCAATTTTATGAAATATTACCTTGGTTTACAGTCTTTTC 1056
 DB 489 GTGATGTTGGAGCATAGGTTGCAATTTTATGAAATATTACCTTGGTTTACAGTCTTTTC 548
 QY 1057 AGACTCATGATAGTAAAGAGCAGCTTGGCAATGATGGAACGAATATTAGGACCCATACAC 1116
 DB 549 AGACTCATGATAGTAAAGAGCAGCTTGGCAATGATGGAACGAATATTAGGACCCATACAC 608
 QY 1117 AACACATGATTCAGAAAAAAGAAAAACGCAAGTATTTTCCACNTAACCCAGTAGATTGG 1176
 DB 609 AACACATGATTCAGAAAAAAGAAAAACGCAAGTATTTTCCACNTAACCCAGTAGATTGG 668
 QY 1177 ATGAACACAGTTCTGCTGGTAGATATGTTAGGAGCGCTGCAAAACCGTTGAAGGAATTTA 1236
 DB 669 ATGAACACAGTTCTGCTGGTAGATATGTTAGGAGCGCTGCAAAACCGTTGAAGGAATTTA 728
 QY 1237 TGTCTTGTGTCATGATGAAGAACATGAGAAACCTGTTTGGCTGTTTGAAGAAATGTTAGAA 1296
 DB 729 TGTCTTGTGTCATGATGAAGAACATGAGAAACCTGTTTGGCTGTTTGAAGAAATGTTAGAA 788
 QY 1297 ATGATCCAACTCAAGAAATTTACCTTGGATGAAGCAATTCGACGATCTCTTTTGTGACTTAT 1356
 DB 789 ATGATCCAACTCAAGAAATTTACCTTGGATGAAGCAATTCGACGATCTCTTTTGTGACTTAT 848
 QY 1357 TAAAAAGAAATGAATGGAATCAGTGTCTTACTATATATCTCTCTAGAGAGATTTAC 1416
 DB 849 TAAAAAGAAATGAATGGAATCAGTGTCTTACTATATATCTCTCTAGAGAGATTTAC 908
 QY 1417 TTAAGACTGTGTCAGTCAACTTAAACATTTCTAATATTTTGTAAACATTTAAATTTTGT 1476
 DB 909 TTAAGACTGTGTCAGTCAACTTAAACATTTCTAATATTTTGTAAACATTTAAATTTTGT 968
 QY 1477 ACAGTTAAGTAAATATTTGATGTTTGTATCAATAGCATAATTAACCTGTTAAGCAAG 1536
 DB 969 ACAGTTAAGTAAATATTTGATGTTTGTATCAATAGCATAATTAACCTGTTAAGCAAG 1028
 QY 1537 TATGCTCTGTAATGCAATTAAGAAATTAATTAATTTTCTTTTGAATTAACCAT 1596
 DB 1029 TATGCTCTGTAATGCAATTAAGAAATTAATTAATTTTCTTTTGAATTAACCAT 1088
 QY 1597 TTTAAATACCTTTCAATATAT - CTTTGTGTCCAGTGAATAATGATGATGATCTTCCCTTT 1655
 DB 1089 TTTAAATACCTTTCAATATATACCTTTGTGTCAGTGAATAATGATGATGATCTTCCCTTT 1148
 QY 1656 TGTACATGGAGGTCACTCTGAAAGTATTTTTTTGAGTAAAAAGAAATCTTGACTA 1712
 DB 1149 TGTACATGGAGGTCACTCTGAAAGTATTTTTTTTGTGAAAAAGAAATCTTGACTA 1205
 XX
 XX RESULT 12
 XX ADI31373
 XX ID ADI31373 standard; cDNA; 1456 BP.
 XX AC ADI31373;
 XX XX 17-JUN-2004 (first entry)
 XX DT
 XX

DE Human CDNA #699.
KW Human; gene; ss; immunological response; immunopathological condition;
KW Crohn's disease; asthma; ulcerative colitis; hyperesoinophilia;
KW irritable bowel syndrome; osteoarthritis; rheumatoid arthritis;
KW acute monocytic leukaemia; antiinflammatory; antiasthmatic; antiulcer;
KW osteopathic; antiarthritic; antirheumatic; cytostatic.
XX
OS Homo sapiens.
XX
XX US6607879-B1.
XX
XX 19-AUG-2003. 98US-00023655.
XX
XX 09-FEB-1998; 98US-00023655.
XX
XX 09-FEB-1998; 98US-00023655.
XX
XX (INCY-) INCYTE CORP.
XX
XX Cocks BG, Stuart SG, Seilhamer JJ;
XX
XX WPI; 2003-895307/82.
XX
XX
XX A composition comprising a plurality of cDNAs, useful for detecting
XX altered expression of genes in an immunological response or for
XX diagnosing and treating an immunopathology, e.g. Crohn's disease, asthma
XX or osteoarthritis.
XX
XX Claim 1; SEQ ID NO 699; 50pp; English.
XX
XX The invention relates to a composition comprising a plurality of cDNAs
XX for detecting the altered expression of genes in an immunological
XX response. The invention also relates to a method of diagnosing or
XX monitoring the treatment of an immunopathological condition in a sample,
XX comprising obtaining nucleic acids from a sample, contacting the nucleic
XX acids of the sample with an array comprising the plurality of cDNAs under
XX conditions to form one or more hybridisation complexes, detecting the
XX hybridisation complexes and comparing the levels of the detected
XX hybridisation complexes with the level of hybridisation complexes
XX detected in a non-diseased sample, where an altered level of the detected
XX hybridisation complexes correlates with the presence of an
XX immunopathological condition. Also disclosed are an expression profile
XX comprising a microarray and a plurality of detectable complexes and a
XX method for identifying a plurality of polynucleotide probes. The cDNAs
XX are useful as hybridisable array elements in a microarray for monitoring
XX the expression of target polynucleotides. The microarray can be used in
XX the diagnosis of an immunopathology, such as Crohn's disease, asthma,
XX ulcerative colitis, hyperesoinophilia, irritable bowel syndrome,
XX osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in
XX identifying agents for the treatment of the diseases. The microarray may
XX also be used in drug discovery and development, toxicological and
XX carcinogenicity studies, forensics or pharmacogenomics. The composition
XX may also be used in purification of a subpopulation of mRNAs, cDNAs or
XX genomic fragments. This sequence represents a human cDNA of the
XX invention. Note: The sequence data for this patent did not form part of
XX the printed specification but was obtained in electronic format directly
XX from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 1456 BP; 480 A; 213 G; 261 G; 502 T; 0 U; 0 Other;
XX
XX Query Match 50.1%; Score 1178.2; DB 11; Length 1456;
XX Best Local Similarity 99.7%; Pred. No. 1.6e-22;
XX Matches 1191; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
XX
XX 1003 TTGGAGCATAGGTTGCATCTTATGAATATACCTTGGTTTCACAGCTTTTCAGACTC 1062
XX
XX 262 TGTAGCGCATAGGTTGCATCTTATGAATATACCTTGGTTTCACAGCTTTTCAGACTC 321
XX
XX 1063 ATGATAGTAAGAGACCTCGCATGATGGAACGAATATTAGGACCCATACACACACA 1122
XX
XX 322 ATGATAGTAAGAGACCTCGCATGATGGAACGAATATTAGGACCCATACACACACA 381

QY 1123 TGATTCAGAAAAACAAGAAAAACGCAAGATATTTTCCACATAAACCCAGCTAGATTGGGATGAAC 1182
DB 382 TGAATTCAGAAAAACAAGAAAAACGCAAGATATTTTCCACATAAACCCAGCTAGATTGGGATGAAC 441
QY 1183 ACAGTTCTGCTGGTAGATATGTTAGGAGACGCTCGAAACCGTTGAAGAAATTTATGCTTTT 1242
DB 442 ACAGTTCTGCTGGTAGATATGTTAGGAGACGCTCGAAACCGTTGAAGAAATTTATGCTTTT 501
QY 1243 GTCATGATGAAGAACATGAGAAAACTGTTTGAACCTGGTTTCGAAGAATGTTAGAAATATGATC 1302
DB 502 GTCATGATGAAGAACATGAGAAAACTGTTTGAACCTGGTTTCGAAGAATGTTAGAAATATGATC 561
QY 1303 CAACTCAAAGAATTAACCTTGGATGAAGCATTCGACGATCCCTTTCTTTGACCTATTATAAAA 1362
DB 562 CAACTCAAAGAATTAACCTTGGATGAAGCATTCGACGATCCCTTTCTTTGACCTATTATAAAA 621
QY 1363 AGAAATGAATGGGAATCAGTGGTCTTACTATATATCTCTCTAGAGAGATTACTTAAGA 1422
DB 622 AGAAATGAATGGGAATCAGTGGTCTTACTATATATCTCTCTAGAGAGATTACTTAAGA 681
QY 1423 CTGTGTCACTCAACTTAAACATTTCTAATATTTTGTAAACATTTAAATATTTTGTACAGTT 1482
DB 682 CTGTGTCACTCAACTTAAACATTTCTAATATTTTGTAAACATTTAAATATTTTGTACAGTT 741
QY 1483 AAGTGTAAATATCTATGTTTGTATCAATAGCATAATTAACCTGTTAAGCAAGTATGCT 1542
DB 742 AAGTGTAAATATCTATGTTTGTATCAATAGCATAATTAACCTGTTAAGCAAGTATGCT 801
QY 1543 CTTGATATGCAATTAGAAAAATTAATAATTTTCTTTTGTAAATTA-CCATTTTAA 1601
DB 802 CTTGATATGCAATTAGAAAAATTAATAATTTTCTTTTGTAAATTA-CCATTTTAA 861
QY 1602 ATACCTTTGAATATCCTTTGTGTCAGTGATAATGTGATGTGATCTTGCCCTTTTGTACA 1661
DB 862 ATACCTTTGAATATCCTTTGTGTCAGTGATAATGTGATGTGATCTTGCCCTTTTGTACA 921
QY 1662 TGGAGGTCACCTCTGAAGTGAATTTTTTTTGTAGTAAAAAGAAATCTTGACTACTTTATAT 1721
DB 922 TGGAGGTCACCTCTGAAGTGAATTTTTTTTGTAGTAAAAAGAAATCTTGACTACTTTATAT 981
QY 1722 CTTAAAGGAATATCTTTATATATCTTCAATTTAGAACTTAACCTTAAAGCTTTTCTTC 1781
DB 982 CTTAAAGGAATATCTTTATATATCTTCAATTTAGAACTTAACCTTAAAGCTTTTCTTC 1041
QY 1782 TGTAAATTTGTAACGGGTGATTAATTAATTAACCTAGATAAGCAGGTACTAGAAAAACAAA 1841
DB 1042 TGTAAATTTGTAACGGGTGATTAATTAATTAACCTAGATAAGCAGGTACTAGAAAAACAAA 1101
QY 1842 CTCAGAAATGTTTACTGTTAGAAATCTATTAATTTTAAAGTGTGATTTCTTTTCAAT 1901
DB 1102 CTCAGAAATGTTTACTGTTAGAAATCTATTAATTTTAAAGTGTGATTTCTTTTCAAT 1161
QY 1902 GGGTGTGTCAGGGTGATAACCAAGCATTCATGAAAGGCGATGAGTTTGTCCATTTGA 1961
DB 1162 GGGTGTGTCAGGGTGATAACCAAGCATTCATGAAAGGCGATGAGTTTGTCCATTTGA 1221
QY 1962 CAGTTTGTAAATAAAACCAATACACATCTTATTTAAGATTAATAATCTAACTGCAAGT 2021
DB 1222 CAGTTTGTAAATAAAACCAATACACATCTTATTTAAGATTAATAATCTAACTGCAAGT 1281
QY 2022 CAGTTTGAATAATGGAATTTCCAAAGTATGTTGGTGAGTCACAGATATAATAATAGAAA 2081
DB 1282 CAGTTTGAATAATGGAATTTCCAAAGTATGTTGGTGAGTCACAGATATAATAATAGAAA 1341
QY 2082 TTCTGATGAGAGGTTTCAGTTTAAATACCAAGTCTTTAGGAGCTTTAAACATTTGCCACG 2141
DB 1342 TTCTGATGAGAGGTTTCAGTTTAAATACCAAGTCTTTAGGAGCTTTAAACATTTGCCACG 1401
QY 2142 ATCTGTTTATCAAAATGACATAAATACGTAACCTTAAAGAAATTAAGTTTATTAAT 2196
DB 1402 ATCTGTTTATCAAAATGACATAAATACGTAACCTTAAAGAAATTAAGTTTATTAAT 1456

RESULT 13
 ID ADI26144
 XX ADI26144 standard; cDNA; 1865 BP.
 AC ADI26144;
 XX
 DT 22-APR-2004 (first entry)
 DE Human cDNA encoding protein that promotes STAT6 activation #55.
 XX
 KW ss; gene; human; signal transducer and activator of transcription 6;
 KW STAT6; immunogen; STAT6 activation; allergy; inflammation; cancer;
 KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
 KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
 KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
 KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.
 XX
 OS Homo sapiens.
 XX
 PN WO2003104277-A2.
 XX
 PD 18-DEC-2003.
 XX
 XX 05-JUN-2003; 2003WO-JP007123.
 XX
 PR 05-JUN-2002; 2002JP-00164257.
 PR 06-JUN-2002; 2002US-0385912P.
 PR 26-DEC-2002; 2002JP-00377326.
 PR 27-DEC-2002; 2002US-0436467P.
 PR 15-MAY-2003; 2003JP-00137505.
 PR 16-MAY-2003; 2003US-0470836P.
 XX
 PA (ASAH) ASAH KASEI KK.
 XX
 PI Sugahara T, Mateuda A, Honda G, Muramatsu S, Ishizawa K;
 XX
 DR WPI; 2004-122214/12.
 XX
 DR P-PSDB; ADI26145.
 XX
 PT New signal transducer and activator of transcription 6 activation
 PT promoting purified protein, for diagnosing and treating disease
 PT associated with activation/inhibition of transcription factor e.g.
 PT diabetes and cancer.
 XX
 PS Claim 4; SEQ ID NO 109; 1368pp; English.
 XX
 CC The invention relates to a purified protein promoting signal transducer
 CC and activator of transcription 6 activation (STAT6). The protein is
 CC useful for the producing an antibody, which involves administering the
 CC protein or its epitope-bearing fragments to a non-human animal as an
 CC antigen. The nucleic acid is useful for diagnosing a disease or
 CC susceptibility to a disease related to expression or activity of the
 CC protein. A transformant expressing the protein is useful for screening
 CC compounds which inhibit or promote STAT6 activation. A transformant
 CC expressing the protein is useful for producing a pharmaceutical
 CC composition. Compositions, antibodies and antisense molecules are useful
 CC for the treating a disease associated with STAT6 activation such as
 CC allergic diseases, inflammation, autoimmune diseases, diabetes,
 CC hyperlipidaemia, infectious disease and cancers. Compositions are useful
 CC for treating disease associated with STAT6 activation and/or prevention
 CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
 CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
 CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
 CC viral hepatitis and AIDS. The protein has efficient promoting STAT6
 CC activity. The protein or nucleic acid is effectively useful for screening
 CC compounds for treating and preventing disease associated with excessive
 CC activation or inhibition of STAT6. The present sequence represents a
 CC human cDNA encoding a protein which promotes STAT6 activation.
 XX
 SQ Sequence 1865 BP; 629 A; 309 C; 381 G; 546 T; 0 U; 0 Other;

Query Match 47.5%; Score 1118.8; DB 12; Length 1865;
 Best Local Similarity 90.2%; Pred. No. 8.3e-211;

	Matches 1191;	Conservative 0;	Mismatches 129;	Indels 0;	Gaps 0;
Qy	400	ATGAAATCGTGGACACATTTGGGTGAAGGAGCGCTTTGGCAAAAGTTGTAGAGTGCATTTGATC	459		
Dd	545	ATGAAATTTGTTGATACATTTAGTGAAGGAGCTTTTGGAAAAGTTGTGGAGTGCATCGATC	604		
Qy	460	ATGGCATGGATGGCATGTCATGTAGCAGTGAATAAATCGTAAAAAATGTAGCCGTTACCGTG	519		
Dd	605	ATAAAGCGGGAGGTAGACATGTAGCAGTAAAAAATAGTTAAAAAATGTGTGATAGATAC	664		
Qy	520	AAGCAGCTCGTTTCAGAAATCCAAGTATTAGAGCAGCTTAAATAGTACTCATCCCAATAGTG	579		
Dd	665	AAGCTGCTCGCTCAGAAATACAGTTCTGGAAACATCTGAATACACAGACCCCAACAGTA	724		
Qy	580	TCITCCGATGTGTCAGATGCTAGATGTTTGTATCATCATGTCGTCATGTTTGTATTTGT	639		
Dd	725	CTTTCCGCTGTGTCAGATGTTGGAATGTTTGGAGCATCATGTCACATTTTGCATTTGTT	784		
Qy	640	TTGAATCTACTGGGACTTGTAGTACTTACGATTTTCAATTAAGAAAAACAGCTTTCTGCCATTTTC	699		
Dd	785	TTGAATCTATTGGGACTTGTAGTACTTACGATTTTCAATTAAGAAAAATGGTTTCTTACCATTTTC	844		
Qy	700	AAATTTGACCATCAGCAGATGGCTATCAGATCTGCCAGTCTCAATTAATTTTACATC	759		
Dd	845	GACTGGATCATATCAGAAAGATGGCATATCAGATATGCAATCTGTGAATTTTTTGCACA	904		
Qy	760	ATAATAAATTAACCCATACAGATCTGAAGCCTGAAAAATATTTTTGTGTAAGTCTGACT	819		
Dd	905	GTAATAGTTGACTCACAACAGCTTAAAGCCTGAAAAACATCTTATTTGTGTCAGTCTGACT	964		
Qy	820	ATGTAAGTCAATATAATTTCTTAAATGAAACCTGTATGAAACGACACTGAAAAACACAGATA	879		
Dd	965	ACACAGAGGGCGTATAATCCCAAAATAAAACCTGTATGAAACGACCTTAAATAATCCAGATA	1024		
Qy	880	TCRAAGTTGTGACTTTGGAGTGCACCTGTATGATGACATCAACATCAGACTTTTGGTGT	939		
Dd	1025	TTAAAGTTGTAGACTTTGGTAGTGCACATATGATGACGAACATCAGACTTATTTGGTAT	1084		
Qy	940	CTACCCGGCCTACAGAGCTCCCGAGGTCAATTTTGGCTTTTAGGTTGGTCTCAGCCCTTGTG	999		
Dd	1085	CTACAAGCATTTAGAGCCTCTGAGTTATTTTAGCCCTAGGGTGGTCCAGCCTTGTG	1144		
Qy	1000	ATGTTTGGAGCATAGTTTGCATTTTATTTGAATATTAACCTTTGGTTTTCAGTCTTTTCAGA	1059		
Dd	1145	ATGTTTGGAGCATAGTTTGGCATTTCTTAATGAATATTAACCTTTGGTTTTCAGTCTTTCCAGA	1204		
Qy	1060	CTCATGATAGTAAGAGCAGCTTGCATGATGGAACGAATATAGGACCCATACACACAC	1119		
Dd	1205	CTCATGATAGTAAGAGCAGCTTGCATGATGGAACGAATATAGGACCCATACACACAC	1264		
Qy	1120	ACATGATTTCAAAAAACAAGAAAAACGCAAGTATTTTCCACATAAACAGCTAGATTGGGATG	1179		
Dd	1265	ACATGATTTCAAAAAACAAGAAAAACGCAAGTATTTTCCACATAAACAGCTAGATTGGGATG	1324		
Qy	1180	AACACAGTTCTGCTGGTAGATATGTTAGGAGACGCTCAAAACCGTTGAAGGAATTTATGC	1239		
Dd	1325	AACACAGTTCTGCTGGTAGATATGTTAGGAGACGCTCAAAACCGTTGAAGGAATTTATGC	1384		
Qy	1240	TTTGTCTATGATGAAGACATGAGAAACCTGTTGACCTGGTTTCGAGAGATTTAGAAATATG	1299		
Dd	1385	TTTGTCTATGATGAAGACATGAGAAACCTGTTGACCTGGTTTCGAGAGATTTAGAAATATG	1444		
Qy	1300	ATCCAACTCAAAAGAAATTAACCTTGGATGAAGCATTTGACAGATCTTTTCTTTCGACTTATTA	1359		
Dd	1445	ATCCAACTCAAAAGAAATTAACCTTGGATGAAGCATTTGACAGATCTTTTCTTTCGACTTATTA	1504		
Qy	1360	AAAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATATCTTCTCTAGAAGAGATTACTTA	1419		
Dd	1505	AAAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATATCTTCTCTAGAAGAGATTACTTA	1564		
Qy	1420	AGACTGTGTCAGTCAACTAAACATCTTAATATTTTGTGTAACATTAATTTTGTGTA	1479		
Dd	1565	AGACTGTGTCAGTCAACTAAACATCTTAATATTTTGTGTAACATTAATTTTGTGTA	1624		


```
Db 1061 GGTCTCAGCCTTGATGTTTGGAGCATAGCTGCATCTTATTGAGTACTACCTTGGGT 1120
Qy 1045 TCACAGTCTTTTCAGACTCATGATAGTAAAGAGCACCTGGCAATGATGGAACGAATATTAG 1104
Db 1121 TCACAGTCTTTTCAGAGCCACCATAGTAAAGAGCACCTGGCAATGATGAGCGGATCTTAG 1180
Qy 1105 GACCCATACCAACACACATGATTCAGAAACAGAAACGCAAGTATTTTCCACCATAC 1164
Db 1181 GACCCATCCAGCATATGATTCAGAAACAGAAACGCAAGTATTTTCCACCATAC 1240
Qy 1165 AGCTAGATTGGATGAACACAGTCTGCTGCTGATATGTTAGGAGACCTGCAAAACCGT 1224
Db 1241 AGCTAGATTGGAGCAGCATAGTTCAGCTGGAGATATGTTAGGAGCCTGCAAGCGT 1300
Qy 1225 TGAAGGAATTTATGCTTTGTCATGATGAAGAACATGAGAACTGTTTGAACCTGTTGCA 1284
Db 1301 TAAAGGAATTTATGCTGCTGTCATGACGAAGAGCATGAGAAGCTGTTTGAACCTGTTGCA 1360
Qy 1285 GAATGTTAGATATGATCAACTCAAGAAATACCTTGGATGAAGCATTTGCAGCATCCTT 1344
Db 1361 GAATGTTGGAGTATGACCCAGCAGAGGATCACCTTGGATGAAGCATTTGCAGCACCTT 1420
Qy 1345 TCTTTGACTTATTAATAAGAAATGA 1370
Db 1421 TCTTTGACTTATTAATAAGAAATGA 1446

RESULT 15
AAD32039 standard; DNA; 21234 BP.
AC AAD32039;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human kinase protein gene.
XX
KW Human; kinase protein; enzyme; cytostatic; osteoplastic; gene expression;
KW colon-moderately differentiated adenocarcinoma; chromosome mapping; gene;
KW drug screening; therapeutic; gene therapy; tissue typing; chromosome 5;
KW bone osteosarcoma; single nucleotide polymorphism; SNP; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace(75..76,A)
FT exon /tag= a
FT /standard_name= "single nucleotide polymorphism"
FT 2007..2059
FT exon /tag= b
FT intron 2060..3118
FT /tag= c
FT exon 3119..3341
FT /tag= d
FT intron 3342..4462
FT /tag= e
FT exon 4463..4553
FT /tag= f
FT intron 4554..4948
FT /tag= g
FT exon 4949..5015
FT /tag= h
FT intron 5016..8054
FT /tag= i
FT variation replace(7980,C)
FT /tag= j
FT /standard_name= "single nucleotide polymorphism"
FT 8055..8171
FT exon /tag= k
FT intron 8172..8258
FT /tag= l
FT exon 8259..8425
FT /tag= m
```

```
FT intron 8426..9007
FT /tag= n
FT variation replace(8571,T)
FT /tag= o
FT /standard_name= "Single nucleotide polymorphism"
FT exon 9008..9102
FT /tag= p
FT intron 9103..9352
FT /tag= q
FT exon 9353..9482
FT /tag= r
FT intron 9483..13437
FT /tag= s
FT variation replace(11257,Y)
FT /tag= t
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(11684,C)
FT /tag= u
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(13312,C)
FT /tag= v
FT exon 13438..13520
FT /tag= w
FT intron 13521..16514
FT /tag= x
FT exon 16515..16594
FT /tag= y
FT intron 16595..17962
FT /tag= z
FT variation replace(17110,T)
FT /tag= aa
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(17451,C)
FT /tag= ab
FT /standard_name= "Single nucleotide polymorphism"
FT exon 17963..18053
FT /tag= ac
FT intron 18054..18139
FT /tag= ad
FT exon 18140..18280
FT /tag= ae
FT variation replace(20766,G)
FT /tag= af
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(20913..20914,T)
FT /tag= ag
FT /standard_name= "Single nucleotide polymorphism"
FT WO200216567-A2.
FT 28-FEB-2002.
FT 24-AUG-2001; 2001WO-US026389.
FT 24-AUG-2000; 2000US-0227470P.
FT 19-MAR-2001; 2001US-00810671.
FT (APPL-) APPLERA CORP.
FT Yan C, Ye J, Ketchum KA, Di Francesco V, Beasley EM;
FT WPI; 2002-269354/31.
FT New human kinase proteins and nucleic acids, useful in drug screening
FT assays, identifying modulators of kinase activity or treating disorders
FT characterized by absence or unwanted expression of the protein.
FT Claim 4; Fig 3; 81pp; English.
XX
XX The invention relates to isolated human kinase proteins and nucleic
XX acids. The nucleic acid and peptide sequences can be used as models for
XX the development of human therapeutic targets, aid in the identification
```

of therapeutic proteins and serve as targets for the development of human therapeutic agents that modulate kinase activity in cells and tissues that express the kinase. The nucleic acids are useful as probes or primers, in constructing recombinant vectors, for expressing antigenic portions of the proteins, chromosome mapping, drug screenings, testing an individual for a genotype, and for gene therapy in patients containing cells that are aberrant in kinase gene expression. The proteins may be used in drug screening assays, in the identification of compounds that modulate, stimulate or inhibit kinase activity, in pharmacogenomic analysis, in treating disorders characterised by an absence or unwanted expression of the protein (bone osteosarcoma, or colon-moderately differentiated adenocarcinoma), and in generating antibodies specific for the peptides. Such antibodies can be used to detect the protein in situ, in vitro, or in cell lysate or supernatant, to isolate and purify the proteins from host cells, pharmacogenomic analysis, tissue typing, and in inhibiting protein function. The present sequence is human kinase protein gene located on chromosome 5

Query Match 46.6%; Score 1097; DB 6; Length 21234;
Best Local Similarity 100.0%; Pred. No. 2.6e-206;
Matches 1097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1228	AGGAATTTATGCTTTGTCATGATGAAGAACATGAGAACTGTTGACCTGTTGCAAGAA	1287
Db	18138	AGGAATTTATGCTTTGTCATGATGAAGAACATGAGAACTGTTGACCTGTTGCAAGAA	18197
Qy	1288	TGTTAGAAATGATCCAACTCAAGAAATACCTTGATGAAGCAATGAGCAATCCTTTCT	1347
Db	18198	TGTTAGAAATGATCCAACTCAAGAAATACCTTGATGAAGCAATGAGCAATCCTTTCT	18257
Qy	1348	TTGACTTATTAAGAAAGAAATGAATGGAAATCAGTGGTCTTACTATATCTTCTAGAA	1407
Db	18258	TTGACTTATTAAGAAAGAAATGAATGGAAATCAGTGGTCTTACTATATCTTCTAGAA	18317
Qy	1408	AGAGATTACTTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTA	1467
Db	18318	AGAGATTACTTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTA	18377
Qy	1468	TTATTTGTACAGTTAAGTGAATAATATGTAATGTTTGTATCAATAGCAATTAATTA	1527
Db	18378	TTATTTGTACAGTTAAGTGAATAATATGTAATGTTTGTATCAATAGCAATTAATTA	18437
Qy	1528	TTAAGCAAGTATGTTCTGTAATGTCATTAGAAATTTAAATTTCTTTCTTTTGAA	1587
Db	18438	TTAAGCAAGTATGTTCTGTAATGTCATTAGAAATTTAAATTTCTTTCTTTTGAA	18497
Qy	1588	ATTACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGAATAATGTAATGATC	1647
Db	18498	ATTACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGAATAATGTAATGATC	18557
Qy	1648	TTGCTTTTGTACATGAGGTCACCTCTGAAGTGAATTTTTTTTGTAGTAAAGGAAATCTT	1707
Db	18558	TTGCTTTTGTACATGAGGTCACCTCTGAAGTGAATTTTTTTTGTAGTAAAGGAAATCTT	18617
Qy	1708	GACTACTTTTATCTTAAAGGAAATTTCTTTATATATCTTCAAAATTTAGAACTTAACCTT	1767
Db	18618	GACTACTTTTATCTTAAAGGAAATTTCTTTATATATCTTCAAAATTTAGAACTTAACCTT	18677
Qy	1768	AAAAAGTTTTTCTCTCTGTAATTTGTTGAACGGGTGATTAATTAACCTCCTAGATAAGCAGGT	1827
Db	18678	AAAAAGTTTTTCTCTCTGTAATTTGTTGAACGGGTGATTAATTAACCTCCTAGATAAGCAGGT	18737
Qy	1828	ACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGAAATCTTATTAATTTTAAGTGTG	1887
Db	18738	ACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGAAATCTTATTAATTTTAAGTGTG	18797
Qy	1888	TATTTCTTTTTCATTTGGGTGATGTCAGGGTGATTAACGACAAATTCATGGAAGGCATGCAG	1947
Db	18798	TATTTCTTTTTCATTTGGGTGATGTCAGGGTGATTAACGACAAATTCATGGAAGGCATGCAG	18857
Qy	1948	TTTGTCCATTTGTGACAGTTTGTGTTTAAATAAAACCATACACACTTTATTTAAGATTAAAA	2007

Db	18858	TTTGTCCATTTGTGACAGTTTGTGTTTAAATAAAACCATACACACTTTATTTAAGATTAAAA	18917
Qy	2008	TCTAACTGGAAGTCAGCTTGGAAAATGGACATTTCCAAAGTATGTTTGGTGAAGTACACAGA	2067
Db	18918	TCTAACTGGAAGTCAGCTTGGAAAATGGACATTTCCAAAGTATGTTTGGTGAAGTACACAGA	18977
Qy	2068	TATAAAATAGAAATTTCTGATGAGAGGTTTTCAGTTTTTAAATACCAAGTCCCTTAGGAGTCT	2127
Db	18978	TATAAAATAGAAATTTCTGATGAGAGGTTTTCAGTTTTTAAATACCAAGTCCCTTAGGAGTCT	19037
Qy	2128	TAACTTTGGCCAGCATCTGTTTATCAAAATGACATAAATACGTAACCTTATAGAAATTAAG	2187
Db	19038	TAACTTTGGCCAGCATCTGTTTATCAAAATGACATAAATACGTAACCTTATAGAAATTAAG	19097
Qy	2188	TTTTATTAATTAGGCAATTTATGCTCTGTGATTAATTTCTTACGGGAGAAAGAGATTTGATTG	2247
Db	19098	TTTTATTAATTAGGCAATTTATGCTCTGTGATTAATTTCTTACGGGAGAAAGAGATTTGATTG	19157
Qy	2248	GAAGCAGTTTGGGAAGAAAGTGTGCTGAAATTTCCAGAAATTTAAATGATTGGTTACAT	2307
Db	19158	GAAGCAGTTTGGGAAGAAAGTGTGCTGAAATTTCCAGAAATTTAAATGATTGGTTACAT	19217
Qy	2308	AACTTTTGTACTTTCAG	2324
Db	19218	AACTTTTGTACTTTCAG	19234

Search completed: March 16, 2005, 09:01:13
Job time : 1257.4 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2005, 07:34:52 ; Search time 7615.01 Seconds
(without alignments)
11766.665 Million cell updates/sec

Title: US-10-801-671-1
Perfect score: 2354
Sequence: 1 gccagctgggtttacttta.....taaaaaaacaaaaaaac 2354

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_ges1: *
9: gb_ges2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1689.4	71.8	2798	3	CR749504	CR749504 Homo sapi
2	1283.2	54.5	1556	3	BC015942	BC015942 Homo sapi
3	1194	50.7	1446	9	AY408252	AY408252 Homo sapi
4	1118.8	47.5	1865	3	AF212224	AF212224 Homo sapi
5	1078.2	45.8	1629	3	AK013974	AK013974 Mus muscu
6	978.2	41.6	1421	9	AY408254	AY408254 Mus muscu
7	937.6	39.8	1446	9	AY408253	AY408253 Pan trogl
8	814.8	34.6	1737	3	CR591789	CR591789 full-leng
9	814.8	34.6	1746	3	CR612115	CR612115 full-leng
10	814.8	34.6	1754	3	CR617472	CR617472 full-leng
11	754.4	32.0	973	5	BUS16568	BUS16568 AGENCOURT
12	734.2	31.2	1689	3	AK076199	AK076199 Mus muscu
13	722	30.7	722	4	BM786882	BM786882 K-EST0065
14	720.8	30.6	966	7	CF110608	CF110608 Shultzomi
15	700.8	29.8	727	1	AA631990	AA631990 np74a05.s
16	700.6	29.8	729	4	BM675318	BM675318 UI-E-EJ0-
17	670.6	28.5	695	5	BQ774390	BQ774390 UI-H-EZ1-
18	666	28.3	1664	3	CR593610	CR593610 full-leng
19	647.6	27.5	1775	3	BC028149	BC028149 Homo sapi
20	644.8	27.4	721	5	BUI184090	BUI184090 AGENCOURT
21	643.2	27.3	659	4	BM681648	BM681648 UI-E-EJ0-
22	641.4	27.2	664	5	BM969890	BM969890 UI-CF-DU1
23	640	27.2	3226	3	BC028573	BC028573 Homo sapi
24	635.2	27.0	643	6	CB127989	CB127989 K-EST0177

25	635.2	27.0	919	4	BG109862	BG109862 602279528
26	632.2	26.9	637	5	BQ729717	BQ729717 UI-E-CK1-
27	628	26.7	1337	7	CO389096	CO389096 AGENCOURT
28	621.8	26.4	849	4	BG916444	BG916444 602814061
29	620.6	26.4	667	1	AJ656344	AJ656344 AJ656344
30	617	26.2	1551	7	CF110517	CF110517 Shultzomi
31	611.6	26.0	957	5	BX374503	BX374503 BX374503
32	609.4	25.9	923	5	BX385193	BX385193 BX385193
33	609	25.9	767	5	BX374502	BX374502 BX374502
34	607	25.8	607	6	CD676847	CD676847 ho07d04.y
35	594.8	25.3	1138	3	CR609176	CR609176 full-leng
36	592.2	25.2	902	1	AL540351	AL540351 AL540351
37	591.6	25.1	950	7	CF411264	CF411264 CH3#072.F
38	591	25.1	624	4	BI494554	BI494554 df11la06.
39	589.8	25.1	593	6	CA314357	CA314357 UI-CF-FNO
40	589.8	25.1	893	6	CB202758	CB202758 AGENCOURT
41	589.2	25.0	950	5	BP383242	BP383242 BP383242
42	573.4	24.4	710	5	BQ210018	BQ210018 UI-R-EPO-
43	570.8	24.2	618	6	CB285812	CB285812 CMD25.A05
44	570	24.2	581	5	BP307701	BP307701 BP307701
45	566	24.0	568	4	BM727179	BM727179 UI-E-EJ0-

ALIGNMENTS

RESULT 1
LOCUS CR749504 2798 bp mRNA linear HTC 19-AUG-2004
DEFINITION Homo sapiens mRNA; cDNA DKFZp686A20267 (from clone DKFZp686A20267).
ACCESSION CR749504
VERSION CR749504.1 GI:51476700
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2798)
AUTHORS Koehrer,K., Beyer,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
Fobo,G., Han,M. and Wiemann,S.
CONSRMT The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMPZ (Biomedical Research Center at the
Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA
sequencing consortium of the German Genome Project. This clone
(DKFZp686A20267) is available at the RZPD Deutsches
Resourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp686A20267
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES

source
1..2798
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686A20267"
/tissue type="cerebellum"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host
DH10B; sites SfiI + SfiIb"
/dev stage="adult"
/note="protein serine threonine kinase Clk4,
differentially spliced"
1..2798
/gene="DKFZp686A20267"
85..246
/gene="DKFZp686A20267"
/codon_start=1
/product="hypothetical protein"

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Smalish, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 1556)

Strausberg, R.

Direct Submission

Submitted (15-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalish, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Teai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAP Plate: 15 Row: d Column: 16

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 10190705

This clone has the following problem: frame shifted.

FEATURES

Location/Qualifiers

1..1556

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3908500"

/tissue_type="Uterus, leiomyosarcoma"

/clone_lib="NIH MGC 71"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

ORIGIN

Query Match 54.5%; Score 1283.2; DB 3; Length 1556;

Best Local Similarity 99.8%; Pred. No. 2.2e-254;

Matches 1285; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 85 GTCATTATTAGAACGAGGTCCTTGAATGACGAGATATTCGGACCGGAGATACGTTG 144

DB 269 GTCATTATTAGAACGAGGTCCTTGAATGACGAGATATTCGGACCGGAGATACGTTG 328

QY 145 ACGAATACAGGAATGACTACTCTGGAAGGATATGTTCTTAGACATTATCACAGACATTG 204

DB 329 ACGAATACAGGAATGACTACTCTGGAAGGATATGTTCTTAGACATTATCACAGACATTG 388

QY 205 AAAGCGGGTATCGAATCCACTGTCAGTAAATCTTCAGTCCGACGAGGAGCAGTCCTA 264

DB 389 AAAGCGGGTATCGAATCCACTGTCAGTAAATCTTCAGTCCGACGAGGAGCAGTCCTA 448

QY 265 AAAGGAAGCGCAATAGACACACTGTTTCAAAGTCATCACTGTCAGTCCGACGAGGAGCAGTCCTA 324

DB 449 AAAGGAAGCGCAATAGACACACTGTTTCAAAGTCATCACTGTCAGTCCGACGAGGAGCAGTCCTA 508

QY 325 AAAGTCCAGGAGTATAGAGATAGATAGAGGAGGTCACCTGATCTGTCTCAATGTGAGAGAGC 384

DB 509 AAAGTCCAGGAGTATAGAGATAGATAGAGGAGGTCACCTGATCTGTCTCAATGTGAGAGAGC 568

QY 385 TTCTAAGAGCAGATATGAAATCGTGGACACTTTTGGGTGAAGGAGCCTTTGCGAAGTTG 444

DB 569 TTCTAAGAGCAGATATGAAATCGTGGACACTTTTGGGTGAAGGAGCCTTTGCGAAGTTG 628

QY 445 TAGAGTGCATTGATCATGGCATGATGTCATGTCATGTCAGTGAAGGAGCCTTTGCGAAGTTG 504

DB 629 TAGAGTGCATTGATCATGGCATGATGTCATGTCATGTCAGTGAAGGAGCCTTTGCGAAGTTG 688

QY 505 TAGGCGGTACCGTGAAGCAGCTCGTTCCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTA 564

DB 689 TAGGCGGTACCGTGAAGCAGCTCGTTCCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTA 748

QY 565 CTGATCCCAATAGTGTCTTCCGATGTCGTCAGATGCTAGATGCTTTCATCATCATGTCGTC 624

DB 749 CTGATCCCAATAGTGTCTTCCGATGTCGTCAGATGCTAGATGCTTTCATCATCATGTCGTC 808

QY 625 ATGTTTGTATGTTTCAACTTACCTGAGCTTAGTACTTACGATTTCAATTAAGAAACAA 684

DB 809 ATGTTTGTATGTTTCAACTTACCTGAGCTTAGTACTTACGATTTCAATTAAGAAACAA 868

QY 685 GCTTTCTGCCATTTCAAAATTGACACATCAGGAGATGCGGTATCAGATCTGCCAGTCAA 744

DB 869 GCTTTCTGCCATTTCAAAATTGACACATCAGGAGATGCGGTATCAGATCTGCCAGTCAA 928

QY 745 TAAATTTTACATCATTAATAAATAAATTAACCCATACAGATCTGAGGCTGAAATATTTGT 804

DB 929 TAAATTTTACATCATTAATAAATAAATTAACCCATACAGATCTGAGGCTGAAATATTTGT 988

QY 805 TTGTGAAGTCTGACTATCTAGTCAAAATAATAATTTCAAAATGAAACGTCATGATGAACGACAC 864

DB 989 TTGTGAAGTCTGACTATCTAGTCAAAATAATAATTTCAAAATGAAACGTCATGATGAACGACAC 1048

QY 865 TGAATAACACAGATATCAAAAGTGTGTGACTTTGGAAGTGCACCGTATGATGAATGACATC 924

DB 1049 TGAATAACACAGATATCAAAAGTGTGTGACTTTGGAAGTGCACCGTATGATGAATGACATC 1108

QY 925 ACAGTACTTTGGTGTCTACCCGGCAGCTACAGAGTCCCGAGGTCATTTTGGCTTTAGGTT 984

DB 1109 ACAGTACTTTGGTGTCTACCCGGCAGCTACAGAGTCCCGAGGTCATTTTGGCTTTAGGTT 1168

QY 985 GGTCTCAGCCTTTGTGATGTTTGGAGCATAGGTTGTCATTTCTTATTCGAATATTAACCTTGGTT 1044

DB 1169 GGTCTCAGCCTTTGTGATGTTTGGAGCATAGGTTGTCATTTCTTATTCGAATATTAACCTTGGTT 1228

QY 1045 TCAGAGTCTTTTCAGACTCATGATAGTAAAGAGCAGCTGCGCAATGATGGAACGAAATATTAG 1104

DB 1229 TCAGAGTCTTTTCAGACTCATGATAGTAAAGAGCAGCTGCGCAATGATGGAACGAAATATTAG 1288

QY 1105 GACCCATACCAACACATGATTTGAGAAACAGAAACCGCAAGTATTTTCCACCAATACC 1164

DB 1289 GACCCATACCAACACATGATTTGAGAAACAGAAACCGCAAGTATTTTCCACCAATACC 1348

QY 1165 AGCTAGATTGGAGTGAACACAGTTCCTGCTGTTAGATATGTTAGGAGAGCGCTGCAACCGT 1224

DB 1349 AGCTAGATTGGAGTGAACACAGTTCCTGCTGTTAGATATGTTAGGAGAGCGCTGCAACCGT 1408

QY 1225 TGAAGGAATTTATGCTTTTGTTCATGATGAAGAACATGAGAAACTGTTTGAACCTGGTTGAA 1284

DB 1409 TGAAGGAATTTATGCTTTTGTTCATGATGAAGAACATGAGAAACTGTTTGAACCTGGTTGAA 1468

QY 1285 GAATGTTTGAATATGATCCAACTCAAGAAATTTACCTTGATGAAGCATTGAGAGATCCTT 1344

Db 1469 GAATGTTAGAAATATGATCCAACTCAAGAANTTACCTTGGATGAAGCATTTGCAGCATCTTT 1529

Qy 1345 TCCTTTGACTTTATTAAAAAAGAAATGAAA 1372
|||||
1529 TCCTTTGACTTTATTAAAAAAGAAATGAAA 1556

RESULT 3
AV408252
LOCUS
DEFINITION Homo sapiens CLK4 gene, VIRTUAL TRANSCRIPT, partial sequence, linear GSS 15-DEC-2003
ACCESSION AY408252
VERSION AY408252.1 GI:39764223
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1446)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1446)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES
source 1..1446
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1446
/gene="CLK4"
/locus_tag="HCM3162"
ORIGIN
Query Match 50.7%; Score 1194; DB 9; Length 1446;
Best Local Similarity 92.8%; Pred. No.5.9e-236;
Matches 1194; Conservative 0; Mismatches -92; Indels 0; Gaps 0;
Qy 85 GTCATTATTAGAACCAAGGTCCTTGAATGACGAGATATTCGGGACCGGAGATACGTTG 144
Db 161 GTCATTATTAGAACCAAGGTCCTTGAATGACGAGATATTCGGGACCGGAGATACGTTG 220
Qy 145 ACGAATACAGGAATCAGCTACTGTGAAGTATGTTCTTAGACATTATCACAGACATTG 204
Db 221 ACGAATACAGGAATGACTACTGTGAAGTATGTTCTTAGACATTATCACAGACATTG 280
Qy 205 AAAGCGGGTATCGAATCCATCGCAGTAAATCTTCAGTCCGACGAGGAGACGATCCTTA 264
Db 281 AAAGCGGGTATCGAATCCATCGCAGTAAATCTTCAGTCCGACGAGGAGACGATCCTTA 340
Qy 265 AAAGGAAGCGGAATAGACACTGTTCAGTCACTAGTCAGTTCGGAAGGCCACCGAAGGA 324
Db 341 AAAGGAAGCGGAATAGACACTGTTCAGTCACTAGTCAGTTCGGAAGGCCACCGAAGGA 400
Qy 325 AAAGATCCAGGAGTATAGAGGATGATGAGGGGTCACTGTATCTGTCAAAGTGGAGACG 384
Db 401 AAAGATCCAGGAGTATAGAGGATGATGAGGGGTCACTGTATCTGTCAAAGTGGAGACG 460
Qy 385 TTCTAAGACGAGATATGAAATCGGTGGAACCTTTTCGGGTGAAGGACGCTTTTGGCAAGTTG 444

AF212224.1	GI:9437514
HTC.	
KEYWORDS	
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 1865)
TITLE	Zhao,M., Song,H., Li,N., Peng,Y., Han,Z. and Chen,Z.
JOURNAL	A novel gene expressed in human bone marrow
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 1865)
TITLE	Zhao,M., Song,H., Li,N., Peng,Y., Han,Z. and Chen,Z.
JOURNAL	Direct Submission
FEATURES	Source
source	1..1865
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/tissue_type="bone marrow"
	154..1515
	/codon_start=1
	/product="CLK4"
	/protein_id="AAF87326.1"
	/db_xref="GI:9437515"
	/translation="MRHSKRTYCPDMDDKDYKWRSSSHKRRKRSHSSAENKRC
	KYNSKWCDSHYLESINSKDYHSRYIDEVNDYTGCEPGHQRDHESRYONHS
	KSSGRSRYSYKXKRRIHSHSRSHGDSIVTLBEGAFKGVCEIDHKAGGRHVA
	LIKVNVDYCEAARSEIQVLEHLNTDPNSFTFCQMLEFHEHGHICVIFELLGLST
	YDITKENGFLPFLRLIKWAYQICKSVNFLSKNKLTHDLKPENILFVQSDYTEAY
	PKIKRDRYLINPDIKVVDGSKATYDDEHSESLVSRHYRAPEVILALGWSQPCDVMS
	IGGLTNYLIGLTVFOETHDSKELAMMERILGPIPOHMIOKTRKRYFHNQLDWBEB
	SSAGRYVRRCCKPLKEPMLCHDEBEKLFPLVRMLLEYDTPQIRITLIDALQHPFDLL
	KKK"
ORIGIN	
Query Match	47.5%; Score 1118.8; DB 3; Length 1865;
Best Local Similarity	90.2%; Pred. No. 2e-220;
Matches 1191; Conservative	0; Mismatches 129; Indels 0; Gaps 0;
Qy	400 ATGAAATCGTGGACACTTTGGTGAAGAGCGCTTTGGCAAAGTTGTAGAGTGCATTGATC 459
Db	545 ATGAAATGTTGTATCTTTAGTGAAGAGCGCTTTGGAAAAGTTGTGGAGTGCATCGATC 604
Qy	460 ATGCGATGGATGGCATCGATGTAGCAGTGAATAATCGTAAAAATGTAGGCGGTACCGTG 519
Db	605 ATAAAGCGGAGGTAGACATGTAGCAGTAAAAATAGTTAAAAATGTGGATAGATACCTGTG 664
Qy	520 AAGCAGCTCGTTCAGAAATCAAGTATAGAGCACTTAAATAGTACTGTATCCCAATAGTG 579
Db	665 AAGCTGCTCGCTCAGAAATCAAGTTCTGGAACATCTGAATACAAGAGCCCAACAGTA 724
Qy	580 TCTTCCGATGTGTCAGATGCTAGAATGGTTTGATCATCATGTCATGTTGTATTGTGTG 639
Db	725 CTTTCCCGCTGTGTCAGATGTGGAATGGTTTGAGCATCATGTGTCACTTTGCATTGTTT 784
Qy	640 TTGAACTACTGGGACTTAGTACTTACCAATTCATTAAAGAAAACAGCTTTCTGCCATTT 699
Db	785 TTGAACCTATTGGGACTTAGTACTTACCACTTCATTAAAGAAAATGGTTTTCTACCAATTC 844
Qy	700 AAATTTGACCAATCAGGCGAGATGCGGTATCAGATCTGCCAGTCAATAAAATTTTTTACATC 759
Db	845 GACTGGATCATATCAGAAAGATGGCATATCAGATATGCAAGTCTGTGAAATTTTTTGCACA 904
Qy	760 ATAATAAATTAACCCATACAGATCTGAAGCCTGAAAAATATTTTGTGTGAAGTCTGACT 819
Db	905 GTAATAAGTTGACTTCACACAGACTTAAAGCCTGAAAACATCTTATTTTGTGCAGTCTGACT 964
Qy	820 ATGTAGTCAAAATATAATCTTAAATGAAACGTGATGAACCGCACTGAAAAACACAGATA 879
Db	965 ACACAGAGGGCGGTATAATCCCAATAAATAAAGTGAATGAACGCACTTAATAAATCCAGATA 1024

TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
REFERENCE	2
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	3
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Iwata, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipipillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	4
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection.
JOURNAL	Nature 409, 685-690 (2001)
MEDLINE	11076861
PUBMED	11076861
REFERENCE	5
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
MEDLINE	6 (bases 1 to 1629)
PUBMED	11076861
REFERENCE	6
AUTHORS	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Horii, F., Inotani, K., Ishii, Y., Itoh, M., Izawa, M., Kusakawa, T., Kato, H., Kawai, J., Kojima, J., Konno, H., Konda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	On Dec 10, 2002 this sequence version replaced gi:12851571. Please visit our web site (http://genome.gsc.riken.jp/) for further details.
FEATURES	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGAGTTAAATTAATCCCGCCCCCCC 3']. cDNA was cleaved with XhoI and SclI. Cloning sites, 5' end: XhoI; 3' end: SclI. Host: SOLR.
FEATURES	Location/Qualifiers

source	1. 1629
	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="C57BL/6J"
	/db_xref="FANTOM_DB:3110002D09"
	/db_xref="taxon:10090"
	/clones="3110002D09"
	/tissue_type="head"
	/clone_lib="RIKEN full-length enriched mouse cDNA library"
	/dev_stage="13 days embryo"
CDS	374..1279
	/note="unnamed protein product; CDC like kinase 4 (MGI:MGI:1098551, GB NM_007714, evidence: BLASTN, 99%, match=984)
	putative"
	/codon_start=1
	/db_xref="GI:26379558"
	/translation="MDGLHVAIKTVKNGVREARSEIQVLEHINSTDPNSVRCVQMLEWFDLHGVCIVFELLGLSTYDFIKENSFLPFQIDHIRMAYQIQSDINFLHNKLTHDLPENILFVKSDVVVKYNSKMRDERTLKNTDIKVDFGSAFYDDEHSTLVSTRHYPAPSVILALGWSQPCDWSIGCILLIYLGFTVFQTHDSKEHLAMMERILGPIPAHMQTKRRKRYFHNNQDDEHSHSAGRYVRRCKPLKPFMLCHDEEHEKULFDLVRML EYDPARRITLDEALQHPFFDLKKK"
ORIGIN	
	Query Match 45.8%; Score 1078.2; DB 3; Length 1629;
	Best Local Similarity 82.5%; Pred. No. 4.9e-212;
	Matches 1351; Conservative 0; Mismatches 183; Indels 103; Gaps 6;
QY	84 AGTCATTATTTAGAACCAAGGTCCTTGAATGAGCGAGATTATCGGACCGGAGATACGTT 143
DB	84 AGTCATTATTTAGAACCAAGGTCCTTGAATGAGCGAGATTATCGGACCGGAGATACATT 143
QY	144 GACGAATACAGGATGACTACTGAGGATATGCTTCTAGACATTTATCAGAGACATT 203
DB	144 GATGAATACAGAAATGACTACTGCGAAGGATATGCTTCAAGACATTACCATAGAGACGTT 203
QY	204 GAAAGCGGATTCGAAATCCACTCCAGTAAATCTTCAGTCCGACGAGGAGAGAGTCCT 263
DB	204 GAAAGCACTTACCGGATCCATTGCGATTAATCTTCAGTCCAGGAGGAGAGAGAGCCCC 263
QY	264 AAAAGGAGCGCAATAGACATGTTCAAGTCATCAGTCACGTTTGAAGAGCCACCGAAGG 323
DB	264 AAGAGAAAGCGTAATAGACCCCTGTCAGTCAAGTCATCAGTCGCAATTCG----- 308
QY	324 AAAAGATCCAGGAGTATAGAGGATGATGAGAGGGGTACCTGATCTGTCAGAGTGGAGAC 383
DB	309 -----
QY	384 GTTCTTAAGAGCAAGATATGAATCGTGACACTTTGGGTGAAGAGCGCTTTGGCAAAAGTT 443
DB	309 -----ATGAATTTGTGGACACTTTTAGTGAAGAGCGCTTTGGCAAAAGTT 352
QY	444 GTAGAGTGCAATGATCATGGCATGGATGGCATGATGATGATGATGATGATGATGATGATGAT 503
DB	353 GTAGAGTGCAATGATCATGGCATGGATGGCATGATGATGATGATGATGATGATGATGATGAT 412
QY	504 GTAGCGGTTACCGTGAAGCAGCTCGTTTCAAGATCCAAATCCAAATCCAAATCCAAATCCAAAT 563
DB	413 GTAGCGGTTACCGGAGGAGCGCTCGTTTCAAGATCCAAATCCAAATCCAAATCCAAATCCAAAT 472
QY	564 ACTGATCCCAATAGTGTCTCCGATGTGTCCAGATGCTAGAGATGCTTGTGATCATCATGCT 623
DB	473 ACTGATCCCAATAGTGTCTCCGATGTGTCCGATGCTAGAGATGCTTGTGATCATCATGCT 532
QY	624 CATGTTGTATGTTTGAACCTACTGGGACTTACTAGTACTTACTAGTACTTACTTAAAGAAAC 683
DB	533 CATGTTGTATGTTTGAACCTACTGGGACTTACTAGTACTTACTTAAAGAAAC 592
QY	684 AGCTTCTGCGATTTCAAAATTCAGATCCAGGAGATGGCGATTCAGATCTGCGAGTCA 743
DB	593 AGTCTTCTGCGATTTCAAAATTCAGATCCAGGAGATGGCGATTCAGATCTGCGAGTCT 652


```

QY 744 ATAAATTTTTCATCATATAATAATTAACCCCATACAGATCTGAAGCCTGAAATAATTTTG 803
DB |||||
QY 653 ATAAATTTTTCATCATATAATAATTAACACACACGACCTAAACCTGAAATAATTTTA 712
DB |||||
QY 804 TTTGTGAAGTCTGACTATGCTCAATATATCTTAAATGAAAGCTGATGAACGCA 863
DB |||||
QY 713 TTTGTGAAGTCTGACTATGCTCAATATATCTTAAATGAAAGCTGATGAACGCA 772
DB |||||
QY 864 CTGAAACACACAGATATCAAGCTTGTGACTTTGGAAGTGCACCTGATGATGAACAT 923
DB |||||
QY 773 TTGAAACACACAGATATCAAGCTTGTGACTTTGGAAGTGCACCTGATGATGAACAT 832
DB |||||
QY 924 CACAGTACTTTGGTCTTACCCGCACTACAGAGCTCCGAGGTCAATTTGGCTTTAGGT 983
DB |||||
QY 833 CATAGTACTTTGGTCTTCCACAGGCACTACAGGCTCCAGAGTCAATTTGGCTTAGGT 892
DB |||||
QY 984 TGGTCTCAGCCTTGTGATGTTTGGAGCATAGTTCATCTTATGAATATACCTTGGT 1043
DB |||||
QY 893 TGGTCTCAGCCTTGTGATGTTTGGAGCATAGTTCATCTTATGAATATACCTTGGG 952
DB |||||
QY 1044 TTCACAGTCTTTCAGACTCATGATAGTAAAGAGCACCTGGCAATGATGAACGAATATTA 1103
DB |||||
QY 953 TTCACAGTCTTTCAGACCCACGATAGTAAAGACCTTGGCAATGATGAACGAATCTTA 1012
DB |||||
QY 1104 GGACCCATACCAACACATGATTCAGAAAACAGAAAACGCAAGTATTTTCCACATTAAC 1163
DB |||||
QY 1013 GGACCCATCCACGACATATGATCCAGAAACAGAAAACGCAAGTATTTTCCACATTAAC 1072
DB |||||
QY 1164 CAGCTAGATTTGGATGAACACAGTCTCTGCTGTAGATATGTTAGAGAGCTGCAACCG 1223
DB |||||
QY 1073 CAGCTAGATTTGGAGCGACATAGTTCAGCTGGAGATATGTTAGAGAGCTGCAACCG 1132
DB |||||
QY 1224 TTGAAGGAATTTATCTTTGTGTCATGATGAAGAAACATGAGAACTGTTTACCTGTTTGA 1283
DB |||||
QY 1133 TTAAGGAATTTATCTTTGTGTCATGACGAGAGCATGAGAGCTGTTTGAACCTGTTTGA 1192
DB |||||
QY 1284 AGAATGTTAGATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCATTTGAGCATCCT 1343
DB |||||
QY 1193 AGAATGTTGGATGATGATCCACGAGCAAGGATCACTTGGATGAAGCATTTGAGCATCCT 1252
DB |||||
QY 1344 TTTCTTGTACTTTATTAAGAAAGAAATGAATGGAATCAGTGTCTTACTATATCTTCTC 1403
DB |||||
QY 1253 TTTCTTGTACTTTATTAAGAAAGAAATGAG-TGGAGTCAAGGGTCTT-CTGTGTACTTCTC 1310
DB |||||
QY 1404 TAGAAGAGATTACTT-AAAGCTGTGTCACTCAACTTAAACATCTTAATTTTGTAAACA 1462
DB |||||
QY 1311 TAGGAGCAGTTACTTCCAGACTGTGTCACTCAACTTAAACCTTCTAATATTTTGTAAACA 1370
DB |||||
QY 1463 TTAATATTTTGTACAGTTAAGTGAATATTTGATGTTTGTATCAATAGCATTAATTA 1522
DB |||||
QY 1371 TTAATATTTTGTACAGTTAAGTGAATATTTGATGTTTGTATCAATAGCATTAATTA 1430
DB |||||
QY 1523 ACTTGTGAAGCAAGTATGCTTGTGATATGATGATGATGATGATGATGATGATGATGATG 1582
DB |||||
QY 1431 CCTTGTGAAGTATGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1490
DB |||||
QY 1583 T-TGAAATTAACATTTTAAATACCTTTGAAATATCTTTGTGTCAGTGAATAACTG 1640
DB |||||
QY 1491 TGAATGTTAACTTTTAAAGGGCTTTTGAATGCCCCCTTTTGTGTCAGTGAATAACTG 1550
DB |||||
QY 1641 ATTGATCTTGGCTTTTGTACATGGAAGGTCACTCTGAAAGTGAATTTTTTTTTT-----GAG 1693
DB |||||
QY 1551 ATTGTCCTCACTTTTGTACATGAAGGTGACTCTGTAAGTGAATTTTTTTTTTCTCTCAAG 1610
DB |||||
QY 1694 TAAAGGAATCTTGAC 1710
DB |||||
QY 1611 TAAAGGAATCTTGAC 1627
DB |||||

```

RESULT 6
AY408254
LOCUS

1421 bp DNA linear GSS 15-DEC-2003

```

DEFINITION Mus musculus CLK4 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY408254
VERSION AY408254.1 GI:39764225
KEYWORDS GSS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1421)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1421)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..1421
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1421
/gene="CLK4"
/locus_tag="HCM3162"
ORIGIN
Query Match 41.6%; Score 978.2; DB 9; Length 1421;
Best Local Similarity 84.1%; Pred. No. 2.2e-191;
Matches 1081; Conservative 0; Mismatches 180; Indels 25; Gaps 1;

```

```

QY 85 GTCATTTATTAGAACGAGTCTTGAATGACGAGATATTATCGGACCGGAGATACGTTG 144
DB |||||
QY 161 GTCATTTATTAGAACGAGTCTTGAATGACGAGATATTATCGGACCGGAGATACATTTG 220
DB |||||
QY 145 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTTATCACAGACATTTG 204
DB |||||
QY 221 ATGAATACAGAAATGACTACTGTGAAGGATATGTTCTTAGACATTTATCACAGACATTTG 280
DB |||||
QY 205 AAAGCGGATCGAATCCACTGCGAGTAAATCTTCACTCCGACGAGGAGAGCAGTCTCTA 264
DB |||||
QY 281 AAAGCATTTACCGGATCCATTTGAGTAAATCTTCACTCCGACGAGGAGAGCAGCCTTA 340
DB |||||
QY 265 AAAGGAGCGCAATAGACATCTGTTCAAGTCACTAGTCACTGTTCAAGAGCAGCAGCAAGGA 324
DB |||||
QY 341 AGAAGAACGTAATAGACCTGTGCAAGTCACTAGTCACTGTTCAAGAGCAGCAGCAAGGA 400
DB |||||
QY 325 AAAGTCCAGCATATAGAGCATGATGAGGAGGTCACTGATCTGTCAAAAGTGGAGAGC 384
DB |||||
QY 401 AAAGTCCAGCATATAGAGCATGATGAGGAGGTCACTGATCTGTCAAAAGTGGAGAGC 460
DB |||||
QY 385 TTCTAAGACGAGATATGAATTCGGACACTTTTGGGTGAAGAGCCTTTTGGCAAGTTG 444
DB |||||
QY 461 TTCTAAGACGAGATATGAATTCGGACACTTTTGGGTGAAGAGCCTTTTGGCAAGTTG 520
DB |||||
QY 445 TAGAGTGCATTTGATCATGGCATGGCATGTCATGTCATGTCATGTCATGTCATGTCATGTC 504
DB |||||
QY 521 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 580
DB |||||
QY 505 TAGGCCGTTACCGTGAAGCAGCTGTTTCAAGAAATCCAAAGTATTAGAGCACTTAATAGTA 564
DB |||||
QY 581 TAGGCCGTTACCGGAGGAGCAGCTGTTTCTGAAATCCAAAGTATTGGAGCACTTGAACACA 640
DB |||||

```

565 CTGATCCCAATAGTGTCTTCGATGTGTCAGATGCTAGATGTTGTTGATCATCATGTC 624
 641 CTGACCCCAACAGTGTCTTCGATGCTCCAGATGCTAGAGTGTGTTGATCATCATGTC 700
 625 ATGTTTGTATGTTGTTGAACCTACTGGGACTTAGTACTTACGATTTTCAATTAAGAAACA 684
 701 ATGTTTGTATGTTGTTGAGCTGCTGGGACTTAGTACTTACGATTTTCAATTAAGAAACA 760
 685 GCTTCTGCCATTTCAATTTGACACCATCAGGAGATGGCGTATCAGATCTGCCAGTCAA 744
 761 GTTTTCTGCCATTTCAATTTGATCAGATCAGGCAATGGC----- 800
 745 TAAATTTTTCATCATATAATTAATTAACCCATACAGATCTGAAGCTCGAAATATTTTGT 804
 801 -----TTTTTACATCATATAATTAATTAACACAGGACTTAACCTGAAATATTTTAT 855
 805 TTGGAAGTCTGACTATGATGATCAATAATAATTTCTAAATGAAAGCTGATGAACGACAC 864
 856 TTGGAAGTCTGACTATGATGATCAATAATAATTTCTAAATGAAAGCTGATGAACGACAC 915
 865 TGAATAACACAGATATCAAGTGTGTTGACTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 924
 916 TGAATAACACAGATATCAAGTGTGTTGACTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 975
 925 ACAGTACTTGTGTTGTTACCCGCACTACAGAGTCTCCGAGGTCAATTTTGGCTTTAGTT 984
 976 ATAGTACTTGTGTTGTTACCCGCACTACAGAGTCTCCGAGGTCAATTTTGGCTTTAGTT 1035
 985 GGTCTCAGCTTGTGATGTTTGGAGCATAGTGTGATTTTATGAAATATTAACCTTTGTT 1044
 1036 GGTCTCAGCTTGTGATGTTTGGAGCATAGTGTGATTTTATGAAATATTAACCTTTGTT 1095
 1045 TCAGAGTCTTTCAGACTCATGATGATGAAGACCTGGCAATGATGAAGCAATATAG 1104
 1096 TCAGAGTCTTTCAGACTCATGATGATGAAGACCTGGCAATGATGAAGCAATATAG 1155
 1105 GACCCATACCAACACATGATTTTCAAGAAACGCAAGTATTTTCAACATATAC 1164
 1156 GACCCATACCAACACATGATTTTCAAGAAACGCAAGTATTTTCAACATATAC 1215
 1165 AGTAGATTTGGATGAACACAGTCTTCTGTTGATGATGTTTGGAGAGCTGCAACCGT 1224
 1216 AGTAGATTTGGAGGAGCATAGTTTCTGTTGATGATGTTTGGAGAGCTGCAACCGT 1275
 1225 TGAAGGATTTATGCTTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1284
 1276 TAAAGGATTTATGCTTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1335
 1285 GAAATTTAGATATGATCAACTCAAGAAATTTACCTTGGATGAAGCATTTGCAGCATCTT 1344
 1336 GAAATTTAGATATGATCAACTCAAGAAATTTACCTTGGATGAAGCATTTGCAGCATCTT 1395
 1345 TCTTTGACTTATTAAGAAAGAAATGA 1370
 1396 TCTTTGACTTATTAAGAAAGAAATGA 1421

RESULT 7
 AY408253
 LOCUS
 DEFINITION Pan troglodytes CLK4 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCESSION AY408253
 VERSION AY408253.1 GI:39764224
 KEYWORDS GSS.
 ORGANISM Pan troglodytes (chimpanzee)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 1 (bases 1 to 1446)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,

Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 Science 302 (5652), 1960-1963 (2003)
 14671302
 2 (bases 1 to 1446)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source
 1. 1446
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 <1..>1446
 /gene="CLK4"
 /locus_tag="HCM3162"
 ORIGIN
 Query Match 39.8%; Score 937.6; DB 9; Length 1446;
 Best Local Similarity 73.1%; Pred. No. 5.3e-183;
 Matches 940; Conservative 0; Mismatches 346; Indels 0; Gaps 0;
 85 GTCATTATTTAGAAGCAAGTCTTGAATGAGCGAGATTTATCGGACCGGAGATACGTTG 144
 161 GTCATTATTTAGAAGCAAGTCTTGAATGAGCGAGATTTATCGGACCGGAGATACGTTG 220
 145 ACGAATACAGGAAGTACTCTGCAAGGATATGTTCTAGACATTTATCAGAGACATTTG 204
 221 ATGAATACAGGAAGTACTCTGCAAGGATATGTTCTAGACATTTATCAGAGACATTTG 280
 205 AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGAGAGGAGAGCAGTCTTA 264
 281 AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGAGAGGAGAGCAGTCTTA 340
 265 AAAGGAAGCGCAATAGACATCTGTTCAAGTCAATGATGATGATGATGATGATGATGATG 324
 341 AAAGGAAGCGCAATAGACATCTGTTCAAGTCAATGATGATGATGATGATGATGATGATG 400
 325 AAAGTCCAGAGTATAGAGGATGATGAGGAGGTCACTGATCTGTCAGAGTGGAGAG 384
 401 AAAGTCCAGAGTATAGAGGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 460
 385 TTCTAAGAGCAAGATATGAAATCGTGGACACTTTTGGTGAAGGAGCCTTTTGGCAAGTTG 444
 461 NNN 520
 445 TAGAGTGCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 504
 521 NNN 580
 505 TAGGCGGTACCGTGAAGCAGCTCGTTTCAGAAATCCAGAGTATTTAGAGCAGTCTTAATAGTA 564
 581 TAGGCGGTACCGTGAAGCAGCTCGTTTCAGAAATCCAGAGTATTTAGAGCAGTCTTAATAGTA 640
 565 CTGATCCCAATAGTGTCTTCCGATGTTGTCAGAGTCTGAGAGTGGTGGTGGTGGTGGTGG 624
 641 CTGATCCCAATAGTGTCTTCCGATGTTGTCAGAGTCTGAGAGTGGTGGTGGTGGTGGTGG 700
 625 ATGTTTGTATTTGTTGTTGAACTACTGGGAGTCTAGTCTTACGATTTTCAATTAAGAAACA 684
 701 ATGTTTGTATTTGTTGTTGAACTACTGGGAGTCTAGTCTTACGATTTTCAATTAAGAAACA 760
 685 GCTTCTGCCATTTCAATTTGACACCATCAGGAGATGGCGTATCAGATCTGCCAGTCAA 744
 761 GCTTCTGCCATTTCAATTTGATCAGATCAGGCAATGGC----- 820

Qy	918	GAACATACAGTACTTTGGTGTCTACCCGGCACTACAGAGCTCCGAGGTCAATTTGGCT	977
Db	1078	GAACATACAGTACTTTGGTGTCTACAGACATATAGACACCTGAGTTATTTAGCC	1137
Qy	978	TTAGTTGGTCTCAGCCTTGATGTTTGGAGCATAGGTGCTTCTTATTGAATATAC	1037
Db	1138	CTAGGGTGGTCCCAACCATGTGATGCTCTGGAGCATAGGATGATCTTTATTGAATATAT	1197
Qy	1038	CTTGGTTTACAGTCTTTACAGTCACTATGATAGTAAAGAGCACTGGCAATGATGAACGA	1097
Db	1198	CTTGGTTTACCGTATTTTCCACACACATAGTATAGGAGCATTTAGCAATGATGAAGAG	1257
Qy	1098	ATATTAGGACCCATACCAACACATGATTCAGAAAAACAAGAAACGCAAGTATTTTCCAC	1157
Db	1258	ATTTTGGACCTCTACCAAAACATATGATACAGAAACCCAGAAACGTAATATTTTCCAC	1317
Qy	1158	CATTAACAGCTAGATTGGATGAACACAGTTCTGCTGTGTAGATATTTAGAGAGCGTGC	1217
Db	1318	CACGATCGATTAGACTGGGATGAACACAGTTCTGCGCGAGATATGTTTCAAGACGCTGT	1377
Qy	1218	AAACCGTTGAAGGAATTTATGCTTTGTCATGATGAAGAAACATGAGAAACTGTTTGACCTG	1277
Db	1378	AAACCTCTGAAGGAATTTATGCTTTCTCAAGATGTTGAACATGAGCGTCTCTTTGACCTC	1437
Qy	1278	GTTTGAAGAAATGTTAGAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCAATTGCAG	1337
Db	1438	ATTCAGAAATGTTGGAGTATGATCCAGCAAAAGAAATTTACTCTCAGAGAAAGCCTTAAAG	1497
Qy	1338	CATCCTTTCTTTGACTTATTTAAAGAAATGAATGGAATCAGTGTCTTACTATATA	1397
Db	1498	CATCCTTTCTTTGACCTTCTGAAGAAAA-----GTATATAGATCTGTAATTTGGACA	1548
Qy	1398	CTTCTCTAGAAGAGATTACTTTAAGACTGTGTCAGTC--AACTAAACATTTCTAAATATTTT	1455
Db	1549	GCTCTCTCGAAGAGA-TCTTACAGACTGTATCAGTCTAATTTTAAATTTTAAGTTATTT	1607
Qy	1456	GTAACATTAATTTATTTGTACAGTTAAGTGAATATTTGATGTTTGTATCAATATAGC	1515
Db	1608	TGTACAGCTTTGTAATTTCTTAACATTTTATTTATTTGCGCATGTTTATTTTGTGGTAA	1667
Qy	1516	ATAATTAATTTGTTAAGCAAGTATGCTTGTATGATGCAATAGAAATTAATTAATTT	1575
Db	1668	TTGGTTTCAATAGTACATAGCTAAGGTAAGTAAGCAATCTTTTTCAGTAATTTGAAGTGT	1727
Qy	1576	TT	1577
Db	1728	TT	1729
RESULT 9			
LOCUS	CR612115	1746 bp	mRNA linear HTC 21-JUL-2004
DEFINITION	full-length cDNA clone CS0DF036YH24 of Fetal brain of Homo sapiens (human).		
ACCESSION	CR612115		
VERSION	CR612115.1	GI:50492922	
KEYWORDS	HTC; CNSLT cDNA.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1. (bases 1 to 1746)		
TITLE	Li, W.B., Gruber, C., Jessee, J., and Polayes, D.		
JOURNAL	Full-length cDNA libraries and normalization		
REMARK	Unpublished		
	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue		
REFERENCE	2. (bases 1 to 1746)		
AUTHORS	Genoscope.		
TITLE	Direct Submission		
JOURNAL	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr		

COMMENT	- Web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life technologies, a division of Invitrogen.		
FEATURES	Location/Qualifiers	source	1. 1746
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="CS0DF036YH24"		
	/tissue_type="Fetal brain"		
	/plasmid="pCMVSPORT_6"		
ORIGIN			
	Query Match	34.6%; Score 814.8; DB 3; Length 1746;	
	Best Local Similarity	73.4%; Pred. No. 1.2e-157; Indels 18; Gaps 4;	
	Matches 1102; Conservative	0; Mismatches 382;	
Qy	84	AGTCATTATTTAGAAAGCAAGGTCTTGAATCAGCGAGATTATCGGGACCGGAGATACGTT	143
Db	237	AGCCATTATTTGGAAAGCAGGTCTATAATCGAAAGATTATCATAGTCGACGTACATT	296
Qy	144	GACGAATACAGAAATGACTACTGTGAAGGATATGTTCTTAGACATTATCACAGACATT	203
Db	297	GATGATACAGAAATGACTACACTCAAGGATGTGAACCTGACACATCGCCAAAGAGACAT	356
Qy	204	GAAAGCGGTATCGAATCCACTGCGAGTAAATTTTCAGTCCGACGAGAGAGAGCTCT	263
Db	357	GAAAGCGGTATCAGAACCATAGTAGCAAGTCTTCTGTAGAGTGGAAAGATGATTAT	416
Qy	264	AAAAGGAAGCCCAATAGACAC-----TGTTCAAGTCACTAGTCACGTTTGAAGAGCCAC	317
Db	417	AAAAGCAACACAGATTTCACACAGTACTTTCACATCGTCTTCACATGGGAAGAGTCAC	476
Qy	318	CGAAGGAAAGATCCAGAGTATAGAGATGATGAGGAGGTGACCTGATCTGTCAAGT	377
Db	477	CGAAGGAAAGAACGAGAGTGTAGAGATGATGAGGAGGTGACCTGATCTGTCAAGT	536
Qy	378	GGAGACGTTCTTAAGAGCAAGATATGAATCTGTGACACTTTGGGTGAAGAGCCCTTTGGC	437
Db	537	GGAGACGTTCTTAAGAGCAAGATATGAATCTGTGACACTTTGGGTGAAGAGCCCTTTGGC	596
Qy	438	AAAGTTGTAGAGTGCATTGATCATGGCATGGATGGCATGTCATGATGAGTGAAGATCGTA	497
Db	597	AAAGTTGTGGAGTGCATCGATCATAAAGCGGAGGTAGACATGTAGCAGTAAAAATAGTT	656
Qy	498	AAAAATGTAGCCGTTACCGTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGAGCATT	557
Db	657	AAAAATGTGGATAGATACCTGTGAAGCTGCTCGCTCAGAAATACAAAGTTCTTGGAACTTG	716
Qy	558	AATAGTACTGATCCCAATAGTGTCTTCGATGTGTCCAGATGCTAGAAATGGTTTGATCAT	617
Db	717	AATACACAGACCCCAACAGTACTTTTCGCTGTGTCCAGATGTTGGAAATGGTTTGAGCAT	776
Qy	618	CATGTCATGTTTGTATTTGTTTGAACACTACTGGGACTTAGTACTTACGATTTCAATTA	677
Db	777	CATGTCATGTTTGTATTTGTTTGAACACTTAGGACTTAGTACTTACGACTTCATTAA	836
Qy	678	GAAACACGCTTTTCGCCATTTTCAAAATGACACATCAGCAGATGGCGCTATCAGATCTGC	737
Db	837	GAAATGGTGTTCCTACCATTTTCGATGATCATATCAGAAAGATGGCATATCAGATATGC	896
Qy	738	CAGTCAATAAATTTTTTACATCAATAAATAAATCCCATACAGATCTGAAGCCTGAAAT	797
Db	897	AAGTCTGCAATTTTTCGACAGTAATTAAGTTGACTCACACAGACTTAAAGCCTGAAAC	956
Qy	798	ATTTTGTGTGAAGTCTGACTATGTTAGTCAAAATATAATTTCTAAATGAACGCTGATGA	857
Db	957	ATCTTATTTGTGAGTCTGACTACACAGAGCGGTATATATCCAAAAATAAAGCTGATGAA	1016
Qy	858	CGCACACTGAAAAACACAGATATCAAGTTTGTGACTTTTGGAAAGTGCACACGTTGATGAT	917

```
Db 1017 CGACCTTAATAATCCAGATATTAAGGTGTAGACTTTGGTAGTGCACATATGATGAC 1076
Qy 918 GAAATCATCAGTACTTTGGTGTCTACCGGCACTACAGAGCTCCGAGTCAATTTGGCT 977
Db 1077 GAACATCATCAGTACTTTGGTGTCTACAGACATTTATAGAGCACTGAAGTATTTTGA 1136
Qy 978 TTAGTGTGGTCTCAGCCTTGTGATGTTTGGAGCATAGGTGTGCAATCTTTATTGAATATTAC 1037
Db 1137 CTAGGGTGTGCTCCCAACCATGTGATGCTGGAGCATAGGATGCAATCTTTATTGAATATTAT 1196
Qy 1038 CTGTGTTTCAAGTCTTTTCAGATCATGATAGTAAAGAGCCTCGGCAATGATGGAACA 1097
Db 1137 CTGTGTTTACCGTATTTTCCACACACGATAGTAAAGAGCACTTTAGCAATGATGGAAGG 1256
Qy 1098 ATATTAGAGCACTACCAACACATGATTTAGAGAAACAGAGAAACGCAAGTATTTCAC 1157
Db 1257 ATCTTGGACCTCTACCAAAACATATGATACAGAAACACAGGAACGTAATATTTCAC 1316
Qy 1158 CATAACCAAGTAGATTGGGATGAACACAGTCTGCTGGTAGATATGTTAGGAGAGCGCTGC 1217
Db 1317 CACGATCGATTAGACTGGGATGAACACAGTCTGCGGCGAGATATGTTTCAAGAGCTGT 1376
Qy 1218 AAACGTTGAAGGAATTTATGCTTTGTCATGATGAAGAACATGAGAAACTGTTTGAACCTG 1277
Db 1377 AAACCTCTGAAGGAATTTATGCTTTCTCAAGATGTTGAACATGAGCGTCTCTTTGACCTC 1436
Qy 1278 GTTCGAAGAAATGTTAGATATGATCCAACTCAAGAAATTAACCTGGATGAAGCAATGCG 1337
Db 1437 ATTCAGAAATGTTGGAGTATGATCCAGCCAAAGAAATTAACCTCAGAGAGCGCTTAAAG 1496
Qy 1338 CATCTCTTCTTGAATTTAAAGAAATGAAATGGAATCACTGCTGCTTACTATATA 1397
Db 1497 CATCTCTTCTTGAATTTGAGAAA-----GTATATAGATCTGTAATTTGGACA 1547
Qy 1398 CTCTCTAGAGAGATTAATTAAGACTGTGTCAGTC--AACTAAACATCTTAATTTT 1455
Db 1548 GCTCTCTCGAAGAGA-TCTTACAGACTGTATCAGTCTAAATTTTAAATTTTAAAGTTATT 1606
Qy 1456 GTAACATTAATTAATTTGTCAGCTTAAGTAAATTAATTTGATGTTTGTATCAATAGC 1515
Db 1607 TGTCAGCTTTGTAATTTCTTAAACATTTTATATTTGCCATGTTTATTTTGTGGTAA 1666
Qy 1516 ATAATTAATCTTTGTAAGCAAGTATGCTCTGATATGATGCAATAGAAAAATTAATTT 1575
Db 1667 TTGGTTTCAATTAAGTACATAGCTAAGGTAATGAACATCTTTTCAGTAATTTGTAAGTAT 1726
Qy 1576 TT 1577
Db 1727 TT 1728

RESULT 10
CR617472
LOCUS
DEFINITION
CR617472 1754 bp mRNA linear HTC 21-JUL-2004
of Homo sapiens (human).
ACCESSION
CR617472.1 GI:50498279
VERSION
HTC; CDS; cDNA.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1754)
AUTHORS
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished
REMARK
Contact : Peng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
REFERENCE
2 (bases 1 to 1754)
AUTHORS
Genoscope.
```

TITLE
JOURNAL

COMMENT

FEATURES

source

ORIGIN

Query Match 34.6%; Score 814.8; DB 3; Length 1754;
Best Local Similarity 73.4%; Pred. No. 1.2e-157;
Matches 1102; Conservative 0; Mismatches 382; Indels 18; Gaps 4;

Direct Submission

Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with NotI and cloned
into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

Location/Qualifiers

1..1754

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clones="CSODI036YH19"

/tissue_type="Placenta Cot 25-normalized"

/plasmid="pCMVSPORT_6"

```
Qy 84 AGTCATATTTAGAGCAAGGTCTTGAATGAGCGAGATTATCGGACCGGAGATACGTT 143
Db 238 AGCCATATTTGGAAAGCAGGTCTATAAATGAGAAAGATTATCATAGTCGACGCTACATT 297
Qy 144 GACCAATACAGGAATGACTACTGTGAGGATATGTTCTAGACATTATCACAGACATT 203
Db 298 GATGAGTACAGAAATGACTACACCAAGGATGTGAACCTCGACATCGCCAAAGACCAT 357
Qy 204 GAAAGCGGTATCGAATCCACTGTCAGTAAATCTTCAGTCCGAGCAGGAGAGAGTCTCT 263
Db 358 GAAAGCGGTATCAGAACCATAGTAGCAAGTCTTCTGTTAGAGTGGAGAGTAGTTAT 417
Qy 264 AAAAGGAAGCCCAATAGACAC-----TGTTCAAGTCAATCAGTCACGTTTCGAAGAGCCAC 317
Db 418 AAAAGCAACACAGAGTTTCAACACAGTACTTTCACATCGTCTGTTCAATCGGAGAGTCCAC 477
Qy 318 CGAAGGAAGAGATCCGAGGTATAGGATGATGAGGAGGCTCACCTGATCTGTCAAAGT 377
Db 478 CGAAGGAAGAGATCCGAGGTATAGGATGATGAGGAGGCTCACCTGATCTGTCAGAGT 537
Qy 378 CGAGACGTTCTTAAGAGCAAGATATGAAATCGTGCACACTTTGGGTGAAGAGCGCTTTGGC 437
Db 538 CGAGACGTTCTTAAGTGCAGATATGAAATCTTGTATCTTTAGGTGAAGAGCTTTTGA 597
Qy 438 AAAGTTGTAGAGTGCATGATCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 497
Db 598 AAAGTTGTAGAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 657
Qy 498 AAAAATGTAGCGCTTACCGTGAAGCAGCTGTTTCAGAAATCCAGATATTAGAGCACCTTA 557
Db 658 AAAAATGTAGCGCTTACCGTGAAGCAGCTGTTTCAGAAATCCAGATATTAGAGCACCTTA 717
Qy 558 AATAGTACTGATCCCAATAGTGTCTTCGATGTGTCCAGATGCTAGAAATGTTTGTATCAT 617
Db 718 AATACACAGAGCCCAACAGTACTTTTCGCTGTGTCCAGATGTTGGATGTTTGGAGCAT 777
Qy 618 CATGCTCATGTTTGTATGTTTGAATCTACTGGGACTTACTGATCTTACGATTTCAATTA 677
Db 778 CATGCTCATGTTTGTATGTTTGAATCTACTGGGACTTACTGATCTTACGATTTCAATTA 837
Qy 678 GAAACACAGCTTTCTGCCATTTCAAAATGACACATCAGGAGAGTGGGTATCAGATCTGTC 737
Db 838 GAAATGTTTCTTACCATTTTCGACTGGATCATATCAGAAAGATGGCATATCAGATATGTC 897
Qy 738 CAGTCAATAAATTTTTCATCATATAAATAAATTAACCCATACAGATCTCAAGCCTGAAAT 797
Db 898 AAGTCTGTGAATTTTTCGACAGATTAAGTTGACTCACACAGACTTAAAGCCTGAAAT 957
Qy 798 ATTTTGTGTGAAGTCTGACTATGATGTAATTAATTTTAAATGAAACGATGAA 857
```

958	ATCTTATTTGTGCGAGTCTGACTACACAGAGGCGTATATATCCCAAAATAAAACGTGATGAA	1017
858	CGCACACTGAAAAACACAGATATCAAAAGTTGTTGACTTTGGAAGTGCACAGTATGATGAT	917
1018	CGCACCTTATAAATCCAGATATTAAGTTGTAGACTTTGAGTGTGCAACATATGATGAC	1077
918	GAACATCACAGTACTTTTGGTGTCTACCCGGCAGCTACAGAGCTCCGAGGTCATTTTGGCT	977
1078	GAACATCACAGTACATTTGGTGTATCTACAAGACATTTATAGAGCACCTGAAGTTATTTTAGCC	1137
978	TTAGGTTGGTCTCAGCCCTCTGTGATGTTTGAGAGCATAGTGTGCATCTCTATTGGAATATTAC	1037
1138	CTAGGTTGGTCCCAACCATGTGATGTCGTGAGCATTAGATGCAATCTTATTGGAATCTAT	1197
1038	CTTGGTTTTCACAGTCTTTTCAGACTCATGATAGTAAAGAGCACCTGGCAATGATGGAAACGA	1097
1198	CTTGGGTTTACCGTATTTTCCAAACACAGATAGTAAAGSAGCATTTAGCAATGATGGAAAGG	1257
1098	ATATTAGGACCCATATACCAACACATGATTCAGAAAAACAGAAAAACCAAGTATTTTCAC	1157
1258	ATTCTTTGGAGCCTCTTCCAAAAACATATGATACAGAAAAACAGGAAACCTAAATATTTTCAC	1317
1158	CATAACCCAGCTAGATTGGGATGAACACAGTTCCTGCTGGTAGATATGTTAGGAGACGCTGC	1217
1318	CAGATCGATTAGACTGGGATGAACACAGTTCCTGCGGCGAGATATGTTTCAAGACGCTGT	1377
1218	AAACCGTTGAAGGAATTTATGCTTTGTCTGATGAAGAAACATGAGAAACCTGTTTGACCTG	1277
1378	AAACCTCTGAAGGAATTTATGCTTTCTCAAGATGTTGAACATGAGCGTCTCTTTGACCTC	1437
1278	GTTTGAAGATGTTAGNATATGATCCAACTCAAGATTAACCTTGGNATGAAGCATTTGCAG	1337
1438	ATTTCAGAAAAATGTTGGAGTAGTATGATCCAGCCAAAGAAATTACTCTCAGAGAAGCCCTTAAAG	1497
1338	CATCCTTTCTTGCATTTATTAATAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATA	1397
1498	CATCCTTTCTTGCATTTCTCAAGAAA-----GTATATAGATCTGTAATTGGACA	1548
1398	CTTCTCTAGAAGAGATTATTAAGACTGTGTCTGATC--AACTAAACATTCTAATATTTTT	1455
1549	GCTCTCTCGAAGAGA-TCTTTACAGACTGTATCAGTCTAAATTTTTTAAATTTTAAAGTTATTT	1607
1456	GTAAACATTAAATATTTTGTACAGTTTAAGTGTAATAATTTGTATGTTTTCATCAATAGC	1515
1608	TGTACAGCTTTGAAATCTTAAACATTTTATATTTGCCATGTTTATTTTGTTTGGGTAAT	1667
1516	ATAATTAACTTGTTTAAGCAAGTAGTGTCTTCGATAATGCATTAGAAAAATTTAAATTAATT	1575
1668	TTGGTTCAITTAAGTACATAGCTAAGGTAATGAACATCTTTTTCAGTAAATTTGTAAGTGAT	1727
1576	TT	1577
1728	TT	1729

RESULT 11	BUS16568	973 bp	linear	EST 12-SEP-2002
LOCUS	BUS16568			
DEFINITION	AGENCOURT 10105484 NIH MGC 134 Mus musculus cDNA clone IMAGE:6513308 5', mRNA sequence.			
ACCESSION	BUS16568			
VERSION	BUS16568.1	GI:22824094		
KEYWORDS	EST.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	1 (bases 1 to 973)			
AUTHORS	NTH-MGC http://mgc.nci.nih.gov/ .			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	Unpublished (1999)			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: ccapbs-r@mail.nih.gov			

```

Tissue Procurement: Dr. David Rowe
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLAM14086 row: m column: 21
High quality sequence stop: 587.
Location/Qualifiers
1. .973
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6513308"
/tissue_type="undifferentiated limb"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 134"
/notes="Vector: pCMV-SPORT6.1, Site_1: EcoRV; Site_2: NotI;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 1.7 kb. Constructed by ResGen, Invitrogen Corp. Note:
this is a NIH MGC Library."
FEATURES
source

```

ORIGIN	Query Match	32.0%;	Score 754.4;	DB 5;	Length 973;
	Best Local Similarity	88.1%;	Pred. No. 3.5e-145;		
	Matches	854;	Conservative 0;	Mismatches 112;	Indels 3; Gaps 3;
QY	478	ATGTAGCAGTGA	AAATCGTTAAAAAATGTTAGGCGGTTACCGTGAAGCAGCTCGTTCAGAAA	537	
DB	6	ATGTAGCAGTGA	AAAAATGTTAAAAAATGTTAGGCGGTTACCGGGAGGCAGCTCGTTCAGAAA	65	
QY	538	TCCCAAGCTATT	TAGAGCAGCTTTAAATATAGTACTGTATCCCAATAGTGTCTTCCGATGTGTCCAGA	597	
DB	66	TCCCAAGTATT	TGGAGCACTTGAACAGCAGCTGACCCCAACAGTGTCTTCCGATGTGTCCAGA	125	
QY	598	TGCTAGAAT	GGTTTGATCATCATCTGGTCACTGTGTTGTAATGTGTTTGAACCTACCTGGGACTTAA	657	
DB	126	TGCTAGAGT	GGTTTGATCATCATCTGGTCACTGTGTTGTAATGTGTTTGAAGTGTCTGGGACTTAA	185	
QY	658	GTA	CTTTACGATTTCAATTAAGAAACAGCTTTCTGCGCATTTTCAAAATGACCAACATCAGGC	717	
DB	186	GTA	CTTATGATTTATTAAGAAAAATAGTTTTCTGCGATTTTCAAAATGATCAATCAGGC	245	
QY	718	AGATGCGGT	ATCAGATCTCGCCAGTCAATAAAATTTTTTATCATCATATAATTAACCCATA	777	
DB	246	AAATGCGCT	ATCAGATCTGCCAGTCTATAAAATTTTTTATCATCATATAATTAACACACA	305	
QY	778	CAGATCTGA	AGCCTGAAAAATATTTTGTGTGAAGTCTGACTATGTAGTCAAAATATAAATT	837	
DB	306	CGSACCT	AAAAACCTGAAAAATATTTTATTTGTGAAGTCTGACTATGTAGTCAAAATATAAATT	365	
QY	838	CTAAAAATG	AAAAACGTGATGAACGCACACTGAAAAACACAGATATCAAAAGTTGTTGACTTTG	897	
DB	366	CTAAAAATG	AAAAACGAGATGAGCGCACATTGAAAAACACAGATATCAAAAGTTGTTGATTTTG	425	
QY	898	GAAGTGC	CAACGTATGATGAACATCACAGTACTTTTGGTGTCTACCCGCGCACTACAGAG	957	
DB	426	GAAGTGC	CAACATATGACGAACATCATATGACTTTTGGTGTCCAACAGGCACTACAGGG	485	
QY	958	CTCCCGAGG	TCATTTTGGCTTTAGGTTGGTCTCAGGCTTTGTGATGTGTTTGGAGCATAGGTT	1017	
DB	486	CTCCAGAGG	TCATTTTGGCTCTAGGTTGGTCTCAGGCTTTGTGATGTGTTTGGAGCATAGGCT	545	
QY	1018	GCATTC	CTTATTTGAATATTACCTTTGGTTTTCACAGTCTTTTCAGACTCATGATAGTAAAGAGC	1077	
DB	546	GCATTC	CTTATTTGAGTACTACCTTTGNGTTTCAGTCTTTTCAGACCCAGTAGTAAAGAGC	605	
QY	1078	ACCTTGC	CAATGATGGAACCGAATATTAGGACCCCATACCACAACACATGATTCAGAAAAACAA	1137	
DB	606	ACCTTGC	CAATGATGGAACCGGATCTTAGGACCCCATCCGACACATATGATCCAGAAAGACAA	665	
QY	1138	GA	AAAAACGCAAGTATTTTTCACCATPAACAGCTAGATTGGGATGAACACAGTTTTCGCTGTA	1197	

Db	666	GAAACGCAAGTATTCACCATACACGCTAGATTGGACGAGCATGTCAGCTGGGA	725
Qy	1198	GATATGTTAGGAGCGCTGCAAAACCGTTGAAGGAATTTATGCTTGTCTCATGATGAAGAAC	1257
Db	726	GATATGTTAGGAGCGCTGCAAGCGTTAAGGAATTTATGCTGTCATGACGAGAGC	785
Qy	1258	ATGAGAACTGTTTGACCTGTTGGAAGAAATGTTAAGATATGATCCAACTCAAGAAATTA	1317
Db	786	ATGAGAAGCTGTTTGACCTGTTTC-AAAAATGTTGGAGTATGATCACCACCGAAGAGTCA	844
Qy	1318	-CCTTGGATGAGCATTCAGCATCTCTTCTTGGACTTATTAATAAAGAAATGAATCGG	1376
Db	845	TCTTTGGATGAGCATTCAGCATCTCTTCTTGGACTTATTAATAAAGAAATGAATCGG	904
Qy	1377	AATCAGTGTCTTACTATATATCTCT-AGAAGAGATTAAGCATGTCAGTCAAA	1435
Db	905	AGCCAGGGGGTCTCGGAACTCCCTTAAGCAGCTTACTCCAGATGGGTGAGCCAA	964
Qy	1436	CTAAACATT 1444	
Db	965	CTAAACCTT 973	
RESULT 12			
AK076199			
LOCUS			
DEFINITION	AK076199	1689 bp mRNA linear HTC 03-APR-2004	
		Mus musculus 18 days pregnant adult female placenta and extra	
		embryonic tissue cDNA, RIKEN full-length enriched library, clone:3830430K12 product:PROTEIN KINASE CLK1 (EC 2.7.1.1-) (CLK)	
ACCESSION	AK076199		
VERSION	AK076199.1	GI:26096662	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			

440 ACAAAGTAAACACAGGAGTCCGCCACCACTTCGCGAGCACCAATTCACACGGGAAGAGTCA 499
 317 CCGAAGGAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCCACCTGATCTGTCAAAG 376
 500 CCGAAGGAAGATGAGGAGGAGTATAGAGGATGATGAGGAGGTCCACCTGATCTGTCAGAG 559
 377 TGGAGAGCTTCTAAGAGCAAGATATGAATCGTGGACACCTTTGGGTGAAGGAGCTTTGG 436
 560 TGGAGAGCTACTAAGTCAAGATATGATATGTTGATACCTTTAGTGAAGGTGCTTTTCGG 619
 437 CAAGTGTAGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 496
 620 AGAAGTGTGGAATGCAATCGCATCAATGATGATGATGATGATGATGATGATGATGATGAT 679
 497 AAAAAAGTACGCGTTCAGCGTGAAGCAGCTCGTTCAAGATCAAGATGATGATGATGATGAT 556
 680 TAAAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 739
 557 AAATAGTACTGATCCCAATAGTGTCTCCGATGTGTCAGATGCTAGAAATGTTGTATCA 616
 740 GTATACAGCAGACCCCATAGTACTTTCCGTTGGTGTCCAGATGTTGGATGTTGTAGCA 799
 617 TCATGTCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 676
 800 TCGAGATCACAATTTGCAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 859
 677 AGAAACAGCTTCTGCAATTTCAAAATTTGACCATCAGGAGGAGTGGGATCAGATCTG 736
 860 GGAACACAGTTTCTGCGGTCTCGAATGTATCATATCAGGAGATGGCATATCAAAATG 919
 737 CCAGTCAATAAATTTTACATCATATAATAAATTAACCCATACAGATCTGAAGCCTGAAAA 796
 920 CAATGCTGTAACCTTTTGCATAGTATATAAATTTGACATACAGATCTGAAGCCTGAAAA 979
 797 TATTTTGTGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 856
 980 CATCTTATTTGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1039
 857 ACGACACTGAAAAACAGATATCAAAATTTGTTGATCTTGGAGTGCACATGATGATGA 916
 1040 ACGTACTATGATTAATCCAGATATTAAGTGTGATGATGATGATGATGATGATGATGATG 1099
 917 TGAACATCAGATGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 976
 1100 TGAACACACAGCATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1159
 977 TTTAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1036
 1160 CTTGGGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1219
 1037 CTTTGGTTTCAAGTCTTTCAGACTCATGATGATGATGATGATGATGATGATGATGATGATG 1096
 1220 TCTTGGATTTACAGTTTTCGATCTCATGATGATGATGATGATGATGATGATGATGATGATG 1279
 1097 AATATTAGGACCCATACCAACATGATGATGATGATGATGATGATGATGATGATGATGATG 1156
 1280 GATTCTTGACCACTTACCAAGGACATGATGATGATGATGATGATGATGATGATGATGATG 1339
 1157 CCATAACAGCTAGTTTGGGATGACACAGTCTGCTGTTGATGATGATGATGATGATGATG 1216
 1340 TCATGATCGATTTAGATTTGGGATGACACAGTCTGCTGTTGATGATGATGATGATGATG 1399
 1217 CAACCGTTGAAGGAATTTATGCTTTTGTATGATGATGATGATGATGATGATGATGATGATG 1276
 1400 TAACTCTGAAGGATTTATGCTATCTCAGATGATGATGATGATGATGATGATGATGATGATG 1459
 1277 GGTTCGAAGATGTTAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1336
 1460 CATTGGGAAAAATGTTGGAGTATGATCCCGCCAAAAGAAATTTACTCTCAAAAGAACCCCTAAA 1519
 1337 GCATCCTTTCTGACTTTTAAAAAAGAA 1366
 1520 GCATCCTTTCTTTTACCCTTTAAAAAGCA 1549

RESULT 13

BM786882

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

BM786882 722 bp mRNA linear EST 05-MAR-2002
 K-EST0065716 S10SNUI Homo sapiens cDNA clone S10SNUI-17-D07 5',
 mRNA sequence.

BM786882
 BM786882.1 GI:19135114
 EST.
 Homo sapiens (human)

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 722)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.

21C Frontier Korean EST Project 2001
 Unpublished (2002)
 Contact: Kim YS

Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr

Plate: 17 row: D column: 07
 High quality sequence stop: 722.
 Location/Qualifiers

1..722
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S10SNUI-17-D07"
 /sex="W"
 /tissue_type="Stomach"
 /cell_type="Lymphoblast-like"
 /cell_line="SNU-1"
 /lab_host="Top10F"
 /clone_lib="S10SNUI"
 /notes="Organ: Stomach; Vector: pME18-FL3; Site: 1: XhoI;
 Site_2: XhoI; The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tobacco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including SfiI
 site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized with Superscript II using SfiI
 oligo-dT primer. After first strand synthesis, RNA was
 degraded by NaOH treatment and cDNA was amplified by PCR
 reaction. The PCR products were digested with SfiI and
 cloned into DraIII- digested pME18S-FL3 vector. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10F' by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library."

ORIGIN

Query Match 30.7%; Score 722; DB 4; Length 722;
 Best Local Similarity 100.0%; Pred. No. 1.7e-138;
 Matches 722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 682 ACAGCTTTCTGCCATTTCAAATTTGACCATCAGGAGATGGCGTATCAGATCTGCCAGT 741
 1 ACAGCTTTCTGCCATTTCAAATTTGACCATCAGGAGATGGCGTATCAGATCTGCCAGT 60
 742 CAATAAATTTTACATCAATAAATTAACCCATCAGATCTGAAGCCTGAAAAATTTT 801
 61 CAATAAATTTTACATCAATAAATTAACCCATCAGATCTGAAGCCTGAAAAATTTT 120
 802 TGTTTGTGAAGTCTGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 861
 121 TGTTTGTGAAGTCTGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180


```

QY 862 CACTGAAAAACACAGATATCAAGTTGTTGACTTTGGAAAGTGCACAGTATGATGATGAAC 921
|
|
|
Db 181 CACTGAAAAACACAGATATCAAGTTGTTGACTTTGGAAAGTGCACAGTATGATGATGAAC 240
|
|
|
QY 922 ATCAGAGTACTTTGTTGTTCTACCCGGCACTACAGAGCTCCCGAGTCAATTTTGGCTTTAG 981
|
|
|
Db 241 ATCAGAGTACTTTGTTGTTCTACCCGGCACTACAGAGCTCCCGAGTCAATTTTGGCTTTAG 300
|
|
|
QY 982 GTTGTCTCAGCTTTGTGATGTTTGGAGCATAGTTGCAATCTTTATTAATATTAACCTTG 1041
|
|
|
Db 301 GTTGTCTCAGCTTTGTGATGTTTGGAGCATAGTTGCAATCTTTATTAATATTAACCTTG 360
|
|
|
QY 1042 GTTTCACAGTCTTTCAGACTCATGATAGTAAGAGCACCTGGCAATGATGGAACGAATAT 1101
|
|
|
Db 361 GTTTCACAGTCTTTCAGACTCATGATAGTAAGAGCACCTGGCAATGATGGAACGAATAT 420
|
|
|
QY 1102 TAGGACCCATACCAACACATGATTTCAGAAACCAAGAAACGCAAGTATTTTCCACCAT 1161
|
|
|
Db 421 TAGGACCCATACCAACACATGATTTCAGAAACCAAGAAACGCAAGTATTTTCCACCAT 480
|
|
|
QY 1162 ACCAGCTAGATTGGATGAACACACAGTCTCTGCTGTGATGATATGTTAGAGAGCGCTGCAAC 1221
|
|
|
Db 481 ACCAGCTAGATTGGATGAACACACAGTCTCTGCTGTGATGATATGTTAGAGAGCGCTGCAAC 540
|
|
|
QY 1222 CGTTGAAGGAATTTATGCTTTGTCTCATGATGAAGAAACATGAGAAACTGTTTGCACCTGGTTC 1281
|
|
|
Db 541 CGTTGAAGGAATTTATGCTTTGTCTCATGATGAAGAAACATGAGAAACTGTTTGCACCTGGTTC 600
|
|
|
QY 1282 GAAGNATGTTAGATATGATCCAACTCAAGNATTAACCTTGGATGAAGCAATGCGAGCATC 1341
|
|
|
Db 601 GAAGNATGTTAGATATGATCCAACTCAAGNATTAACCTTGGATGAAGCAATGCGAGCATC 660
|
|
|
QY 1342 CTTTCTTTGACTTATTAAGAAAGAAATGAATGGAATCAGTGGTCTTACTATATATCTTC 1401
|
|
|
Db 661 CTTTCTTTGACTTATTAAGAAAGAAATGAATGGAATCAGTGGTCTTACTATATATCTTC 720
|
|
|
QY 1402 TC 1403
|
|
Db 721 TC 722

RESULT 14
CF110608
LOCUS Shultz.M.A., Zhang, L., Gu, Y.-Z., Baker, G.L., Panmuchi, M.V.,
DEFINITION Shultz.M.A., Gurske, W.A., Morin, S.G., Jovanovich, S.B.,
Rattus norvegicus cDNA clone Contig3421 5', mRNA sequence.
ACCESSION CF110608
VERSION CF110608.1 GI:33166640
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
-ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 966)
AUTHORS Shultz.M.A., Zhang, L., Gu, Y.-Z., Baker, G.L., Panmuchi, M.V.,
Padua, A.M., Gurske, W.A., Morin, S.G., Jovanovich, S.B.,
Flopper, C.G. and Buckpitt, A.R.
TITLE Gene expression analysis in response to lung toxicants: I.
JOURNAL Sequencing and microarray development
COMMENT Am. J. Respir. Cell Mol. Biol. 30 (3), 296-310 (2004)
Contact: Shultz MA
Dept. of Molecular Biosciences, School of Veterinary Medicine
University of California, Davis
1311 Haring Hall, One Shields Avenue, Davis, CA 95616, USA
Tel: 530 752 0793
Fax: 530 752 4698
Email: mashultz@ucdavis.edu
Average Phred score is 20 or better. All poor quality data (Phred <
20) and vector/linker sequence has been removed.
High quality sequence stop: 966.
Location/Qualifiers

```

```

source 1..966
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="Contig3421"
/sex="male"
/tissue_type="airway or parenchyma"
/dev_stage="adult"
/clone_lib="Rat lung airway and parenchyma cDNA libraries"
/notes="Organ: lung; Vector: pGEM-11zf(-); Site 1: Eco RI;
Site 2: Not I; mRNA was isolated from microdissected rat
lung airways and parenchyma tissues."

```

ORIGIN

```

Query Match      30.6%; Score 720.8; DB 7; Length 966;
Best Local Similarity 86.0%; Pred. No. 3.1e-138;
Matches 822; Conservative 0; Mismatches 133; Indels 2; Gaps 2;

QY 749 TTTTTCATCATTAATAAATTAACCCATACAGATCTGAAGCTGAAATATTTTGTGT 808
|
|
|
Db 1 TTTTTCATCATTAATAAATTAACACACACGCGATCTAAACCTGAAATATTTTATCGT 60
|
|
|
QY 809 GAAGCTGACTATGCTCAATATATTTCTAAATGAAACGTGATGAACGACACTGAA 868
|
|
|
Db 61 AAAGTCTGACTATGTGGTCAAAATCAATTTAAATGAAACGAGATGAGCGACGTTGAA 120
|
|
|
QY 869 AAACACAGATATCAAAAGTTGTTGACTTTGGAAGTGCACAGTATGATGAACATCAG 928
|
|
|
Db 121 AAACACAGATATCAAAAGTTGTTGACTTTGGAAGTGCACATATGACGATGAACATCATAG 180
|
|
|
QY 929 TACTTTGTGTCTACCCGGCACTACAGAGCTCCGAGGTCAATTTTGGCTTTAGTGTGTC 988
|
|
|
Db 181 TACTTTGTGTCTACCAAGACACTACAGGCTCCAGAGTCAATTTTGGCTTTAGTGTGTC 240
|
|
|
QY 989 TCAGCTTGTGATGTTTGGAGCATAGTTGCAATCTTATGAAATATTAACCTTGTTCAC 1048
|
|
|
Db 241 TCAGCTTGTGATGTTTGGAGCATAGGCTGCAATCTTATGAGTATTAACCTTGGATTCCAC 300
|
|
|
QY 1049 AGTCTTTCAGACTCATGATAGTAAAGAGCACCTGGCAATGATGGAACCAATATTAGGACC 1108
|
|
|
Db 301 AGTCTTTCAGACTCATGATAGTAAAGAACACCTGGCAATGATGGAAGCGGATCTCGGGCC 360
|
|
|
QY 1109 CATACCAACACACATGATTTCAAGAAAACAAAGAAAACGCAAGTATTTTCCACATAACCAAGCT 1168
|
|
|
Db 361 CATCCAGCACACATGATCCAAAAGACAAAGAAACGCAAGTATTTTCCACATAACCAAGCT 420
|
|
|
QY 1169 AGATTGGAGTGAACACAGTCTCTGCTGTGATGATATGTTAGGAGACGCTGCAAAACCGTTGAA 1228
|
|
|
Db 421 AGATTGGAGTGAACATAGTTTCAGCTGGGAGATAGCTTTAGGAGACGCTGCAAGCCGTTAAA 480
|
|
|
QY 1229 GGAATTTATGCTTTGTCTATGATGAAGAACATGAGAAACCTGTTGACCTGGTTTGAAGAAT 1288
|
|
|
Db 481 GGAATTTATGCTTTGTCTACGATGAAGAGCATGAGAAAGCTGTTTACCTGGTTTGAAGAAT 540
|
|
|
QY 1289 GTTAGAATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCAATTTGACAGCATCTCTTTCT 1348
|
|
|
Db 541 GTTGGAGTATGATCCCAACAAAGAAAGATTAACCTTGGATGAAGCTTTGACAGCAACCTTTCT 600
|
|
|
QY 1349 TGACTTATTAAGAAAAGAAATGAAGATGGAATCAGTGGTCTTACTATATATCTTCTAGAA 1408
|
|
|
Db 601 TGACTTATTAAGAAAAGAAATG-AATGGAAATCAGTGGTCTTCTCTATATATCTTCTTAGGA 659
|
|
|
QY 1409 GAGATTTAC-TTAAGACTGTGTCAGTCACTAAACATTTCTAATATTTTGTAAACATTAAA 1467
|
|
|
Db 660 GCAATTTACTTTAAGACTGTGTCAGTCACTAAACATTTCTAATATTTTGTAAACATTAAA 719
|
|
|
QY 1468 TTATTTGTACAGTTAAGTGAATATTTGATGTTTGTATCAATAGCATATTAATTAACCTTG 1527
|
|
|
Db 720 TTATTTGTACAGTTAAGTGAATATTTGATGTTTGTATCAATAGCATATTAATTAACCTTG 779
|
|
|
QY 1528 TTAAGCAAGTATGCTTCTTGATATGCAATAGAAAATTAATTAATTTTCTTTTGTGAA 1587
|
|
|

```


780 TTAAGTATGTTGATGAATGAATCAATATAGTAAATGAATTTTCCTTGTGATGTTA 839

1588 ATTACCACTTTTAATATACCTTTGAATATCCTTTGTGTCAGTGCATTAATGATGATGATC 1647

840 ATTGCCATTTTAAAGGCTTTACAGTCCCTTTGTGTCAGTGGTAAACGTTGATGATGTC 899

1648 TTGCTTTTGTACATGAGGAGTCACTCTGAAAGTGAATTTTTTTTGGTAAAGGAAA 1703

900 TCACTTTTGTACATTAAGTTAACTTTGAAGTGGTTTTTTTTTCCCTCAAGTAA 955

RESULT 15
AA631990/c

LOCUS
DEFINITION
727 bp mRNA linear EST 28-OCT-1997
np74a05.s1 NCI CGAP Br2 Homo sapiens cDNA clone IMAGE:1132016 3'
similar to SW:KCLK_MOUSE P22518 PROTEIN KINASE CLK ; mRNA
sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AA631990
AA631990.1 GI:2555404
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 727)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 2117 Std Error: 0.00
Seq primer: -40m3 fwd. Et from Amersham
High quality sequence stop: 475.
Location/Qualifiers
1. 727
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1132016"
/sex="female, pooled"
/tissue type="breast"
/lab host="DH10B"
/clone.lib="NCI CGAP Br2"
/note="vector: pVT3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from pooled bulk
breast tumor tissue, and was then primed with a Not I -
oligo(3T) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pVT3
vector. This library is the normalized version of
NCI CGAP Br1.1. Library was constructed by Bento Soares
and M. Fatima Bonaldo."

ORIGIN

Query Match 29.8%; Score 700.8; DB 1; Length 727;
Best Local Similarity 99.5%; Pred. No. 4e-134;
Matches 724; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

988 CTCAGCCTTTGTGATGTTTGGAGCATAGCTTCATTCTTTATTCATATTAATTAACCTTGGTTTCA 1047

726 CTCAGCCTTTGTGATGTTTGGAGCATAGCTTCATTCTTTATTCATATTAATTAACCTTGGTTTCA 668

1048 CAGTCTTTGAGTATGATAGTAAAGAGCAGCTTGGCAATGATGGAACGAATATTAGGAC 1107

667 CAGTC-TTCAGACTCATGATAGTAAAGAGCAGCTTGGCAATGATGGAACGAATATTAGGAC 609

1108 CCATACCAACAACACATGATTCAGAAAAACAAGAAAAACGCAAGTATTTTCCACATAACCAAGC 1167

608 CCATACCAACAACACATGATTCAGAAAAACAAGAAAAACGCAAGTATTTTCCACATAACCAAGC 549

1168 TAGATTGGGATGAACACAGTTTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACGGTTGA 1227

548 TAGATTGGGATGAACACAGTTTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACGGTTGA 489

1228 AGGAATTTATGCTTTGTCATGATGAAGAAACATGAGAAACTGTTTGACCTGCTTGAAGAA 1287

488 AGGAATTTATGCTTTGTCATGATGAAGAAACATGAGAAACTGTTTGACCTGCTTGAAGAA 429

1388 TGTTAGAAATATGATCCAACTCAAGAAATACCTTTGGATGAAGCAATGGCAGCATCTTTCT 1347

428 TGTTAGAAATATGATCCAACTCAAGAAATACCTTTGGATGAAGCAATGGCAGCATCTTTCT 369

1348 TTGACTTTATTAAAAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATATCTTCTAGA 1407

368 TTGACTTTATTAAAAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATATCTTCTAGA 309

1408 AGAATTTACTTAAGACTGTGTCAGTCACTCACTCAATTAATTTTGTAAACATTTAAA 1467

308 AGAATTTACTTAAGACTGTGTCAGTCACTCACTCAATTAATTTTGTAAACATTTAAA 249

1468 TTATTTTGTACAGTTAAAGTGAATATTTGATGTTTGTATCAATAGCATTAATTAACCTTG 1527

248 TTATTTTGTACAGTTAAAGTGAATATTTGATGTTTGTATCAATAGCATTAATTAACCTTG 189

1528 TTAAGCAAGTATGCTCTTGATAATGCAATAGAAAAATTAATAATTTTCTTTTGA 1587

188 TTAAGCAAGTATGCTCTTGATAATGCAATAGAAAAATTAATAATTTTCTTTTGA 129

1588 ATTACCACTTTTAAATACCTTTTGAATATCCTTTGTGTCAGTGCATTAATGATGATGATC 1647

128 ATTACCACTTTTAAATACCTTTTGAATATCCTTTGTGTCAGTGCATTAATGATGATGATC 69

1648 TTGCTTTTGTACATGAGGAGTCACTCTGAAAGTGAATTTTTTTTGGTAAAGGAAATCTT 1707

68 TTGCTTTTGTACATGAGGAGTCACTCTGAAAGTGAATTTTTTTTGGTAAAGGAAATCTT 9

1708 GACTACTT 1715

8 GACTACTT 1

Search completed: March 16, 2005, 18:40:31
Job time : 7627.01 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2005, 07:43:58 ; Search time 393.184 Seconds
(without alignments)
9796.410 Million cell updates/sec

Title: US-10-801-671-1
Perfect score: 2354
Sequence: 1 gccagctgggtacttta.....taaaaaaacaaaaaaac 2354

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2354	100.0	2354	3	US-09-810-671-1
2	2354	100.0	2354	4	US-10-109-854-1
3	2354	100.0	2354	4	US-10-339-656-1
4	2052.6	87.2	2446	2	US-09-016-000-9
5	1178.2	50.1	1456	4	US-09-023-655-699
6	1097.4	46.6	1549	4	US-09-905-999-26
7	1097	46.6	21234	3	US-09-810-671-3
8	1097	46.6	21234	4	US-10-109-854-3
9	1097	46.6	21234	4	US-10-339-656-3
10	640.6	27.2	2254	4	US-09-919-039-238
11	481.2	20.4	1538	4	US-09-905-999-22
12	454.8	19.3	1787	4	US-09-905-999-24
13	432.4	18.4	1762	4	US-09-016-434-1439
14	430.8	18.3	1763	4	US-09-949-016-2648
15	430.8	18.3	1763	4	US-09-949-016-2649
16	430.8	18.3	1788	4	US-09-976-594-313
17	263	11.2	263	4	US-09-016-434-310
18	229.2	9.7	913	4	US-09-016-434-712
19	203.4	8.6	492	4	US-09-621-976-3125
20	185.4	7.9	479	4	US-09-621-976-3124
21	123.8	5.3	378	1	US-08-700-575-2
22	95.2	4.0	2061	2	US-08-835-170-1
23	95.2	4.0	2061	3	US-09-359-257-1
24	95.2	4.0	2061	3	US-09-371-674-1
25	95.2	4.0	2327	2	US-08-835-170-3
26	95.2	4.0	2327	3	US-09-359-257-3
27	95.2	4.0	2327	3	US-09-371-674-3

28	89	3.8	621	4	US-09-248-796A-4385	Sequence 4385, Ap
29	84.8	3.6	308	4	US-09-621-976-9631	Sequence 9631, Ap
30	81.4	3.5	260	4	US-09-313-294A-2299	Sequence 2299, Ap
31	75.6	3.2	3723	4	US-09-949-016-3590	Sequence 3590, Ap
32	75.6	3.2	71251	4	US-09-949-016-15332	Sequence 15332, A
33	68.8	2.9	282	4	US-09-313-294A-1168	Sequence 1168, Ap
34	67.2	2.9	2424	4	US-09-614-221A-518	Sequence 518, App
35	66.4	2.8	466	4	US-09-513-999C-11373	Sequence 11373, A
36	66	2.8	362	4	US-09-513-999C-14437	Sequence 14437, A
37	63.6	2.7	187169	4	US-09-949-016-12776	Sequence 12776, A
38	63.6	2.7	191569	4	US-09-949-016-15940	Sequence 15940, A
39	61.8	2.6	25882	4	US-09-949-016-14390	Sequence 14390, A
40	61.8	2.6	25882	4	US-09-806-708B-22	Sequence 22, Appl
41	60.6	2.6	1141	4	US-09-949-016-13504	Sequence 13504, A
42	60.6	2.6	59519	4	US-09-949-016-12378	Sequence 12378, A
43	60.6	2.6	119153	4	US-08-487-826B-13	Sequence 13, Appl
44	60.4	2.6	19124	2	US-09-949-016-17296	Sequence 17296, A
45	60.4	2.6	251672	4	US-09-949-016-17296	Sequence 17296, A

ALIGNMENTS

RESULT 1
US-09-810-671-1
; Sequence 1, Application US/09810671
; Patent No. 6455291
; GENERAL INFORMATION:
; APPLICANT: YAN Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO00758
; CURRENT APPLICATION NUMBER: US/09/810,671
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Human
US-09-810-671-1

Query Match	100.0%;	Score 2354;	DB 3;	Length 2354;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2354;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GCAGCTGGGGTTACTTTTAAAAAATGCTCCATGTGATCCCTCTTGAAGTTGCACT	60	
Db	1	GCAGCTGGGGTTACTTTTAAAAAATGCTCCATGTGATCCCTCTTGAAGTTGCACT	60	
Qy	61	CTGTTGAAGAGGACACTCCAGTCAATTTTAAAGGAGGTCCTTGAATGAGCGAG	120	
Db	61	CTGTTGAAGAGGACACTCCAGTCAATTTTAAAGGAGGTCCTTGAATGAGCGAG	120	
Qy	121	ATTATCGGACCGGAGATACCTTGACGAATACAGGAATGACTCTGTGAAGGATATGTC	180	
Db	121	ATTATCGGACCGGAGATACCTTGACGAATACAGGAATGACTCTGTGAAGGATATGTC	180	
Qy	181	CTAGACATTATTCACAGAGACATTGAAAGCGGATATCGAATCCACTGCGAGTAATCTTCAG	240	
Db	181	CTAGACATTATTCACAGAGACATTGAAAGCGGATATCGAATCCACTGCGAGTAATCTTCAG	240	
Qy	241	TCCGACAGGAGGAGGAGTCTTAAAGGAGCGCAATGACACTGTTCAAGTCATCAGT	300	
Db	241	TCCGACAGGAGGAGGAGTCTTAAAGGAGCGCAATGACACTGTTCAAGTCATCAGT	300	
Qy	301	CACGTTTGAAGAGGACCGAAGGAAAGATCCAGGATATAGAGGATCATAGGAGGGTC	360	
Db	301	CACGTTTGAAGAGGACCGAAGGAAAGATCCAGGATATAGAGGATCATAGGAGGGTC	360	
Qy	361	ACCTGATCTGTCAGAGTGGAGACGTTCTTAAGAGCAAGATATGAAATCGTGACACTTGG	420	
Db	361	ACCTGATCTGTCAGAGTGGAGACGTTCTTAAGAGCAAGATATGAAATCGTGACACTTGG	420	


```
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-109-854-1

Query Match      100.0%; Score 2354; DB 4; Length 2354;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCAGCTGGGTTACTTTAAAAAACATGCTCATGTGCTGCTCTTGAAGCTTCGCAC 60
DB 1 GCCAGCTGGGTTACTTTAAAAAACATGCTCATGTGCTGCTCTTGAAGCTTCGCAC 60

QY 61 CTGTTGAAGAGACACTCATCCAGTCAATTTTGAAGCAGAGTCCCTTGAATGAGCGAG 120
DB 61 CTGTTGAAGAGACACTCATCCAGTCAATTTTGAAGCAGAGTCCCTTGAATGAGCGAG 120

QY 121 ATTATCGGACCGGAGATACGTTGACGAATACAGGAATGACTACTGTGAGGATATGTT 180
DB 121 ATTATCGGACCGGAGATACGTTGACGAATACAGGAATGACTACTGTGAGGATATGTT 180

QY 181 CTAGACATATACAGAGACATTTGAAGCGGATATCGAATCCACTGCGATTAATCTTCAG 240
DB 181 CTAGACATATACAGAGACATTTGAAGCGGATATCGAATCCACTGCGATTAATCTTCAG 240

QY 241 TCCGACGAGGAGAGACGCTCTTAAGAGCGCAATAGACACTGTTCACTGATCATAGT 300
DB 241 TCCGACGAGGAGAGAGCAGTCTTAAGAGCGCAATAGACACTGTTCACTGATCATAGT 300

QY 301 CACGTTTCCAAGAGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATGAGGAGGCT 360
DB 301 CACGTTTCCAAGAGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATGAGGAGGCT 360

QY 361 ACCTGATCTGCAAGTGGAGACGTTCTAAGAGCAAGATATGAATTCGTGGACATTTGG 420
DB 361 ACCTGATCTGCAAGTGGAGACGTTCTAAGAGCAAGATATGAATTCGTGGACATTTGG 420

QY 421 GTGAAGGAGCCTTTGGCAAGGTTGTAGAGTGCATTTGATCATGCGCATGTCATGCGATG 480
DB 421 GTGAAGGAGCCTTTGGCAAGGTTGTAGAGTGCATTTGATCATGCGCATGTCATGCGATG 480

QY 481 TAGCAGTGAATAATCGTAAATAATGAGCCGTTACCGTGAAGCAGCTCGTTTCAGAAATCC 540
DB 481 TAGCAGTGAATAATCGTAAATAATGAGCCGTTACCGTGAAGCAGCTCGTTTCAGAAATCC 540

QY 541 AAGTATTAGACACTTAATAATAGTACTGATCCCAATAGTGTCTTCGATGTCAGATGC 600
DB 541 AAGTATTAGACACTTAATAATAGTACTGATCCCAATAGTGTCTTCGATGTCAGATGC 600

QY 601 TAGAATGTTTGCATCATGTCATGTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 660
DB 601 TAGAATGTTTGCATCATGTCATGTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 660

QY 661 CTTACGATTTCAATAAGAAACAGCTTCTGCCATTTCAATTCACACATCAGGCGAGA 720
DB 661 CTTACGATTTCAATAAGAAACAGCTTCTGCCATTTCAATTCACACATCAGGCGAGA 720

QY 721 TGGCGTATCAGATCTGCCAGTCAATAAATTTTTTACATCAATAAATTAACCCATACAG 780
DB 721 TGGCGTATCAGATCTGCCAGTCAATAAATTTTTTACATCAATAAATTAACCCATACAG 780

QY 781 ATCTGAAGCCTGAAATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 840
DB 781 ATCTGAAGCCTGAAATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 840
```

```
QY 841 AAATGAAACGTGATGAACGCACTAAGAAAAACACAGATATCAAGATTTGTTGATTTGGAA 900
DB 841 AAATGAAACGTGATGAACGCACTAAGAAAAACACAGATATCAAGATTTGTTGATTTGGAA 900

QY 901 GTGCAACGTATGATGATGAACATCACTACAGTACTTTGGTGTCTACCCGGCACTACAGACTC 960
DB 901 GTGCAACGTATGATGATGAACATCACTACAGTACTTTGGTGTCTACCCGGCACTACAGACTC 960

QY 961 CCGAGGTCAATTTTGGCTTTAGTGTCTGACGCTTTGATGTTTGGAGCATAGGTTGCA 1020
DB 961 CCGAGGTCAATTTTGGCTTTAGTGTCTGACGCTTTGATGTTTGGAGCATAGGTTGCA 1020

QY 1021 TTTCTATTGAATATTAACCTTTGGTTCACAGCTTTTTCAGACTCATCATGATAGTAAAGACAC 1080
DB 1021 TTTCTATTGAATATTAACCTTTGGTTCACAGCTTTTTCAGACTCATCATGATAGTAAAGACAC 1080

QY 1081 TGGCAATGATGGAACGAATATTTAGGACCCATACCAACACATGATTCAGAAAAACAGAA 1140
DB 1081 TGGCAATGATGGAACGAATATTTAGGACCCATACCAACACATGATTCAGAAAAACAGAA 1140

QY 1141 AACGCAAGTATTTTCCACATTAACAGCTAGATTGGGATGAACACACAGTTCCTGCTGTAGAT 1200
DB 1141 AACGCAAGTATTTTCCACATTAACAGCTAGATTGGGATGAACACACAGTTCCTGCTGTAGAT 1200

QY 1201 ATGTTAGAGACGCTGCAACCGTTGAAGGAATTTATGCTTTGTCATGATGAAGAACATG 1260
DB 1201 ATGTTAGAGACGCTGCAACCGTTGAAGGAATTTATGCTTTGTCATGATGAAGAACATG 1260

QY 1261 AGAACTGTTTGAACCTGCTGCAAGAAATTTAGAAATATGATCCAACTCAAGAAATTAACCT 1320
DB 1261 AGAACTGTTTGAACCTGCTGCAAGAAATTTAGAAATATGATCCAACTCAAGAAATTAACCT 1320

QY 1321 TGGATGAAGACGCTGCAAGCACTCTTTTGTGCTTAAAGAAAGAAATGAAATGGAATC 1380
DB 1321 TGGATGAAGACGCTGCAAGCACTCTTTTGTGCTTAAAGAAAGAAATGAAATGGAATC 1380

QY 1381 AGTGTCTTACTATATATCTCTAGAAAGAGATTAATTAAGACTGTGTGCTGCTCAACTAAA 1440
DB 1381 AGTGTCTTACTATATATCTCTAGAAAGAGATTAATTAAGACTGTGTGCTGCTCAACTAAA 1440

QY 1441 CATCTAATATTTTGTAAACATTAATTAATTTGTACAGTAACTGATTAATTTGTTATG 1500
DB 1441 CATCTAATATTTTGTAAACATTAATTAATTTGTACAGTAACTGATTAATTTGTTATG 1500

QY 1501 TTTTGTATCAATAGCATAAATTAATCTGTAAAGAGTATGCTTGTGATTAATTCATAGAA 1560
DB 1501 TTTTGTATCAATAGCATAAATTAATCTGTAAAGAGTATGCTTGTGATTAATTCATAGAA 1560

QY 1561 AAATTAATAATTAATTTTCTTTTGAATTTACCAATTTTAAATACCTTTGAAATATCCTTT 1620
DB 1561 AAATTAATAATTAATTTTCTTTTGAATTTACCAATTTTAAATACCTTTGAAATATCCTTT 1620

QY 1621 TGTGTCAGTGAATAATGATGATGATCTTGTGCTTTGTGATCATGAGAGTCACTCTGAAAT 1680
DB 1621 TGTGTCAGTGAATAATGATGATGATCTTGTGCTTTGTGATCATGAGAGTCACTCTGAAAT 1680

QY 1681 GATTTTTTTTTCAGTAAAGGAAATCTTTCACACTTATATTTTAAAGGAAATTTCTTTTA 1740
DB 1681 GATTTTTTTTTCAGTAAAGGAAATCTTTCACACTTATATTTTAAAGGAAATTTCTTTTA 1740

QY 1741 TATACCTTCAAAATTTAGAACTTTAACTTTTAAAGTTTTTCTTCTGTAAATTTGTTGAAACGGGTG 1800
DB 1741 TATACCTTCAAAATTTAGAACTTTAACTTTTAAAGTTTTTCTTCTGTAAATTTGTTGAAACGGGTG 1800

QY 1801 ATTATTAATACTCTAGATAAGCGGTAATAAGAAACCAAACTCAGAAAAATGTTTACTGTT 1860
DB 1801 ATTATTAATACTCTAGATAAGCGGTAATAAGAAACCAAACTCAGAAAAATGTTTACTGTT 1860

QY 1861 TAGAATTTCTAATAATTTTAAAGTGTGTTGTTTCTTTTCAATTTGGGTTGTTGTTGTTGTT 1920
DB 1861 TAGAATTTCTAATAATTTTAAAGTGTGTTGTTTCTTTTCAATTTGGGTTGTTGTTGTTGTT 1920

QY 1921 ACCAGACATTCATGGAAGGCAATGACAGTTTGTCCATTTGTGACAGTTTGTGTTTAAATAAAC 1980
```

```

Db 1921 ACCAGACATTCATGGAAGGCGATGCGATTTGTCCATTTGTGACAGTTTCTTTAATAAACC 1980
QY 1981 ACATACACACATTTTAAAGATTAAATCTAACTGGAAGTCAGCTTGGAAAATGGACAT 2040
Db 1981 ACATACACACATTTTAAAGATTAAATCTAACTGGAAGTCAGCTTGGAAAATGGACAT 2040
QY 2041 TTCCAAGTATGTTGGTGAGTCACAGATATAAAATAGAAAATCTGATGAGAGGTTTCAG 2100
Db 2041 TTCCAAGTATGTTGGTGAGTCACAGATATAAAATAGAAAATCTGATGAGAGGTTTCAG 2100
QY 2101 TTTTAAATACCAAGTCCTTAGAGTCCTTAAATAGAAAATCTGATGAGAGGTTTCAG 2160
Db 2101 TTTTAAATACCAAGTCCTTAGAGTCCTTAAATAGAAAATCTGATGAGAGGTTTCAG 2160
QY 2161 TAAATACCTTAACCTATTAAGATTAAAGTTTAAATAGGCAATTTATGCTGTGATAAT 2220
Db 2161 TAAATACCTTAACCTATTAAGATTAAAGTTTAAATAGGCAATTTATGCTGTGATAAT 2220
QY 2221 TCTTACGGGAGAAAGAGGATTTGATTGGAAGCAGTTTGGGAAGAAAGTCTGCTGAAAT 2280
Db 2221 TCTTACGGGAGAAAGAGGATTTGATTGGAAGCAGTTTGGGAAGAAAGTCTGCTGAAAT 2280
QY 2281 TTCCAGAAATTTAATGATTGTTACATATAAATTTTGTGACTTTCAGAAAAAATAAATA 2340
Db 2281 TTCCAGAAATTTAATGATTGTTACATATAAATTTTGTGACTTTCAGAAAAAATAAATA 2340
QY 2341 AACAAAAAATAAAC 2354
Db 2341 AACAAAAAATAAAC 2354

```

```

RESULT 3
US-10-339-656-1
; Sequence 1, Application US/10339656
; Patent No. 673978
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758DIV2
; CURRENT APPLICATION NUMBER: US/10/339,656
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/109,854
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-339-656-1

```

```

Query Match 100.0%; Score 2354; DB 4; Length 2354;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCAGCTGGGTTACTTTAAAAAACATGCTCCATGTCATCCCTCTTGAAGCTTCGCACT 60
Db 1 GCCAGCTGGGTTACTTTAAAAAACATGCTCCATGTCATCCCTCTTGAAGCTTCGCACT 60
QY 61 CTGTTGAAGAGGACACTCATCCAGTCATTTTAAAGAGCAAGGTCTTGAATAGCGAG 120
Db 61 CTGTTGAAGAGGACACTCATCCAGTCATTTTAAAGAGCAAGGTCTTGAATAGCGAG 120
QY 121 ATTATCGGACCGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGATATGTC 180
Db 121 ATTATCGGACCGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGATATGTC 180

```

```

QY 181 CTAGACATTTATCACAGAGACATTTGAAAGCGGATTCGAATCCACTGCAAGTAAATCTTCAG 240
Db 181 CTAGACATTTATCACAGAGACATTTGAAAGCGGATTCGAATCCACTGCAAGTAAATCTTCAG 240
QY 241 TCCGAGAGAGAGAGAGAGTCTCTTAAAGAGGAGCCAAATAGACACTGTTCAAGTCAATCAGT 300
Db 241 TCCGAGAGAGAGAGAGTCTCTTAAAGAGGAGCCAAATAGACACTGTTCAAGTCAATCAGT 300
QY 301 CACGTTCCGAGAGAGAGAGATCCGGAAGGAAAGATCCAGAGTATAGAGATGATGAGAGGCTC 360
Db 301 CACGTTCCGAGAGAGAGAGATCCGGAAGGAAAGATCCAGAGTATAGAGATGATGAGAGGCTC 360
QY 361 ACCTGATCTGTCAAAGTGGAGACGTTCTTAAGAGCAAGATATGAAATCGTGGACACTTTGG 420
Db 361 ACCTGATCTGTCAAAGTGGAGACGTTCTTAAGAGCAAGATATGAAATCGTGGACACTTTGG 420
QY 421 GTGAAGAGAGCCTTTGGCAAAAGTGTAGAGTGCATGATGCGCATGGCATGGCATGCGATG 480
Db 421 GTGAAGAGAGCCTTTGGCAAAAGTGTAGAGTGCATGATGCGCATGGCATGGCATGCGATG 480
QY 481 TAGCAGTGAAATCGTAAAAAATGTTAGGCGGTTTACCGTGAAGCAGCTCGTTTCAGAAATCC 540
Db 481 TAGCAGTGAAATCGTAAAAAATGTTAGGCGGTTTACCGTGAAGCAGCTCGTTTCAGAAATCC 540
QY 541 AAGTATTAGAGCACTTAAATAGTACTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGC 600
Db 541 AAGTATTAGAGCACTTAAATAGTACTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGC 600
QY 601 TAGAATGGTTTGTGATCATGCTGATGTTGTGATGTTGTGTTGAACTACTGGGACTTAGTA 660
Db 601 TAGAATGGTTTGTGATCATGCTGATGTTGTGATGTTGTGTTGAACTACTGGGACTTAGTA 660
QY 661 CTTACGATTTCAATTAAGAGAAACAGCTTTCTGCCATTTTCAAAATGACCACTACAGGAGCA 720
Db 661 CTTACGATTTCAATTAAGAGAAACAGCTTTCTGCCATTTTCAAAATGACCACTACAGGAGCA 720
QY 721 TGGCGTATCAGATCGCCAGTCAATAAATTTTACATCATATAAATTAATTAATTAATTAAT 780
Db 721 TGGCGTATCAGATCGCCAGTCAATAAATTTTACATCATATAAATTAATTAATTAATTAAT 780
QY 781 ATCTGAAGCCTGAAAATATTTTGTGTAAGTCTGACTATGTAAGTCAAAATTAATTTCTA 840
Db 781 ATCTGAAGCCTGAAAATATTTTGTGTAAGTCTGACTATGTAAGTCAAAATTAATTTCTA 840
QY 841 AAATGAAGAGTGTGATGAACGACACCTGAAAGAACACAGATATCAAAAGTCTGTTGACCTT 900
Db 841 AAATGAAGAGTGTGATGAACGACACCTGAAAGAACACAGATATCAAAAGTCTGTTGACCTT 900
QY 901 GTGCAACGATGATGATGAACATCACAGTACTTTGGTGTCTACCCGGCAGCTACAGAGCTC 960
Db 901 GTGCAACGATGATGATGAACATCACAGTACTTTGGTGTCTACCCGGCAGCTACAGAGCTC 960
QY 961 CCGAGGTCATTTGGCTTTAGGTTGGTCTCAGCCTTGTGATGTTGGAGCATAGGTTGCA 1020
Db 961 CCGAGGTCATTTGGCTTTAGGTTGGTCTCAGCCTTGTGATGTTGGAGCATAGGTTGCA 1020
QY 1021 TTTCTATTGAATATTTACCTTGGTTTTCAGCTTTTCAAGTCTCATGATAGTAAAGAGCACC 1080
Db 1021 TTTCTATTGAATATTTACCTTGGTTTTCAGCTTTTCAAGTCTCATGATAGTAAAGAGCACC 1080
QY 1081 TGGCAATGATGGAAGCAATATTAGGACCCATACCAACACACATGATTCAGAAAAACAAGAA 1140
Db 1081 TGGCAATGATGGAAGCAATATTAGGACCCATACCAACACACATGATTCAGAAAAACAAGAA 1140
QY 1141 AACGCAAGTATTTTCAACCAAGTGTAGATGGGATGGAACAGCTTCTCTGCTGATAGT 1200
Db 1141 AACGCAAGTATTTTCAACCAAGTGTAGATGGGATGGAACAGCTTCTCTGCTGATAGT 1200
QY 1201 ATGTTAGGAGAGCGTGCAGAAACCGTTGAAGGAATTTATGCTTTGTGTCATGATGAAGAACATG 1260
Db 1201 ATGTTAGGAGAGCGTGCAGAAACCGTTGAAGGAATTTATGCTTTGTGTCATGATGAAGAACATG 1260

```

QY 1261 AGAACTCTTTGACCTGGTTCGAGAAATGTTAGAAATATGATCCAACTCAAGAAATTAACCT 1320
Db 1261 AGAACTCTTTGACCTGGTTCGAGAAATGTTAGAAATATGATCCAACTCAAGAAATTAACCT 1320
QY 1321 TGGATGAAGCATTCGAGCATCTTTCTTTGACTTATTAAGAAAGAAATGAATGGGATC 1380
Db 1321 TGGATGAAGCATTCGAGCATCTTTCTTTGACTTATTAAGAAAGAAATGAATGGGATC 1380
QY 1381 AGTGTCTTACTATATCTTCTCTAGAGAGATTAAGAGATGTTGTCAGTCAACTAAA 1440
Db 1381 AGTGTCTTACTATATCTTCTCTAGAGAGATTAAGAGATGTTGTCAGTCAACTAAA 1440
QY 1441 CATCTCAATATTTTGTAAAATTAATATTTTGTACAGTAAGTGAATATTTGTATG 1500
Db 1441 CATCTCAATATTTTGTAAAATTAATATTTTGTACAGTAAGTGAATATTTGTATG 1500
QY 1501 TTTTGTATCAATAGCATATTAATCTTTGTAAGCAAGTATGCTTTGATAATGCCATTAGAA 1560
Db 1501 TTTTGTATCAATAGCATATTAATCTTTGTAAGCAAGTATGCTTTGATAATGCCATTAGAA 1560
QY 1561 AAATTAATAATTTTCTTTTGAATTAACCAATTTTAAATACCTTTTGAATATCCTT 1620
Db 1561 AAATTAATAATTTTCTTTTGAATTAACCAATTTTAAATACCTTTTGAATATCCTT 1620
QY 1621 TGTGTCAGTGAATAATGTGATGATCTTGGCTTTTGTACATGAGAGTCACTCTGAAAT 1680
Db 1621 TGTGTCAGTGAATAATGTGATGATCTTGGCTTTTGTACATGAGAGTCACTCTGAAAT 1680
QY 1681 GATTTTTTTCAGTAAAAGGAAATCTTCACACTTATATTTCTTAAAGGAATATCTTTA 1740
Db 1681 GATTTTTTTCAGTAAAAGGAAATCTTCACACTTATATTTCTTAAAGGAATATCTTTA 1740
QY 1741 TATACTTCAAAATTTAGAACTTTAACTTTTAAAGTTTCTTCTGTAATTTGTTGAACGGGTG 1800
Db 1741 TATACTTCAAAATTTAGAACTTTAACTTTTAAAGTTTCTTCTGTAATTTGTTGAACGGGTG 1800
QY 1801 ATTAATTAATCTCTAGATAAGCAGTACTAGAAACCAAACTCAGAAAATGTTTACTGT 1860
Db 1801 ATTAATTAATCTCTAGATAAGCAGTACTAGAAACCAAACTCAGAAAATGTTTACTGT 1860
QY 1861 TAGAATCTATTAATTTTAAAGTGTGTTATCTTTTCAATTTGGGTGATGTCAGGTGATA 1920
Db 1861 TAGAATCTATTAATTTTAAAGTGTGTTATCTTTTCAATTTGGGTGATGTCAGGTGATA 1920
QY 1921 ACCAGACATTCATGAAAGGATGAGTTTGTCCATTTGTGACAGTTTCTTAAATAAACCC 1980
Db 1921 ACCAGACATTCATGAAAGGATGAGTTTGTCCATTTGTGACAGTTTCTTAAATAAACCC 1980
QY 1981 ACATACACACTTTATTTAAGATTAATAATCTAACTGGAAGTCAGCTTGGAAAATGGACAT 2040
Db 1981 ACATACACACTTTATTTAAGATTAATAATCTAACTGGAAGTCAGCTTGGAAAATGGACAT 2040
QY 2041 TTCCAGTATGTTGGTGCAGTCACAGATATAAATAAGAAATTTCTGATGAGAGTTTCAG 2100
Db 2041 TTCCAGTATGTTGGTGCAGTCACAGATATAAATAAGAAATTTCTGATGAGAGTTTCAG 2100
QY 2101 TTTTAAATACCAAGTCCTTAGAGCTTAAATGCGGAGCATCTGTTTATCAATGACA 2160
Db 2101 TTTTAAATACCAAGTCCTTAGAGCTTAAATGCGGAGCATCTGTTTATCAATGACA 2160
QY 2161 TAAATACGTAACCTTATTAAGAAATTAAGTTTATTAATAGGCAATTTATGCTGTGATAAT 2220
Db 2161 TAAATACGTAACCTTATTAAGAAATTAAGTTTATTAATAGGCAATTTATGCTGTGATAAT 2220
QY 2221 TCTTACGGGAGAAAGAGATTTGATTGGAAGCAGTTTGGGAAGAAAGTGTCTGTAAT 2280
Db 2221 TCTTACGGGAGAAAGAGATTTGATTGGAAGCAGTTTGGGAAGAAAGTGTCTGTAAT 2280
QY 2281 TTCCAGAAATTTAATTTGATTTGATACATAAACTTTTGTGACTTTCAGAAAAAATAAAAA 2340
Db 2281 TTCCAGAAATTTAATTTGATTTGATACATAAACTTTTGTGACTTTCAGAAAAAATAAAAA 2340
QY 2341 AACAAAAAATAAAC 2354

Db 2341 AACAAAAAATAAAC 2354

RESULT 4

US-09-016-000-9
; Sequence 9, Application US/09016000
; Patent No. 5962232
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Bandman, Olga
; APPLICANT: Akerman, Ingrid B.
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; TITLE OF INVENTION: PROTEIN KINASE MOLECULES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,000
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0465 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2446 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: NEUTIFMT01
; CLONE: 339963
US-09-016-000-9

Query Match 87.2%; Score 2052.6; DB 2; Length 2446;

Best Local Similarity 95.8%; Pred. No. 0;
Matches 2155; Conservative 0; Mismatches 4; Indels 90; Gaps 1;

QY 85 GTCATTATTTAGAACGAAGTCTTGAATGAGCGAGATTAATCGGACCGGAGATACGTTG 144
Db 288 GTCATTATTTAGAACGAAGTCTTGAATGAGCGAGATTAATCGGACCGGAGATACGTTG 347
QY 145 ACGAATACAGGAATGACTTCTGTGAAGGATATGTTCTTAGACATTAATCAGAGACATTG 204
Db 348 ACGAATACAGGAATGACTTCTGTGAAGGATATGTTCTTAGACATTAATCAGAGACATTG 407
QY 205 AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGAGCAGGAGAGAGCAGTCTCTA 264
Db 408 AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGAGCAGGAGAGAGCAGTCTCTA 467
QY 265 AAAGGAAGCGCAATAGACACTGTTCAAGTCAATCAGTCACTGTTTCAAGAGAGCCACCGAAGGA 324

622	AGAAATGAATGGGAATCAGTGGTCTTATATATATCTTCTCTAGAGAGATTACTTAAGA	681
1423	CTGTGTCAAGTCAACTAAACAATCTTAATATATTTTGTGTAAACAATTAATATTTGTGTACAGTT	1482
682	CTGTGTCAAGTCAACTAAACAATCTTAATATATTTTGTGTAAACAATTAATATTTGTGTACAGTT	741
1483	AAAGTGTAAATATTTGTATGTTTTGTGTATCAATAGACATTAATTAACCTGTGTAAAGCAAGTATGGT	1542
742	AAAGTGTAAATATTTGTATGTTTTGTGTATCAATAGACATTAATTAACCTGTGTAAAGCAAGTATGGT	801
1543	CTTGATAATGCAATTTAGAAAAAATTAATATTTTCTTTTTCGAAATTA-CCATTTTTTAA	1601
802	CTTGATAATGCAATTTAGAAAAAATTAATATTTTCTTTTTCGAAATTAATTAATTTTAA	861
1602	ATACCTTTGAAATATCTTTTGTGTCCAGTGATAAATGTGATGATCTGTGCTTTTGTGACA	1661
862	ATACCTTTGAAATATCTTTTGTGTCCAGTGATAAATGTGATGATCTGTGCTTTTGTGACA	921
1662	TGAGGTCACTCTGAAGTGATTTTTTTTGTGAGTAAAGGAAATCTTGACCTATTTATATT	1721
922	TGAGGTCACTCTGAAGTGATTTTTTTTGTGAGTAAAGGAAATCTTGACCTATTTATATT	981
1722	CTTAAAGGAATATCTTTTATATATCTTTCAAATTTAGAACTTAACCTTTAAAGTTTTTCTTC	1781
982	CTTAAAGGAATATCTTTTATATATCTTTCAAATTTAGAACTTAACCTTTAAAGTTTTTCTTC	1041
1782	TGTAATTTGTTGAAACGGGTGATTTATTTAACTCTAGATAGCAGGTACTAGAAACCAAAA	1841
1042	TGTAATTTGTTGAAACGGGTGATTTATTTAACTCTAGATAGCAGGTACTAGAAACCAAAA	1101
1842	CTCAGAAAAATGTTTACTGTTAGAAATCTATTAATAATTTTAAAGTGTGTATTTCTTTTCATT	1901
1102	CTCAGAAAAATGTTTACTGTTAGAAATCTATTAATAATTTTAAAGTGTGTATTTCTTTTCATT	1161
1902	GGGTGATGTCAGGGTGATTAACACAGACATTCATCGAAAGGCATCGAGTTTCTCCATTGTGA	1961
1162	GGGTGATGTCAGGGTGATTAACACAGACATTCATCGAAAGGCATCGAGTTTCTCCATTGTGA	1221
1962	CAGTTTGTGTTTAAATAAAACCAATACATACACATTTTATTTAAAGATTAATACTAACTGGAAAGT	2021
1222	CAGTTTGTGTTTAAATAAAACCAATACATACACATTTTATTTAAAGATTAATACTAACTGGAAAGT	1281
2022	CAGCTTGGAAAAATGGACATTTCCAAGTATGTTTTGGTGGAGTACACAGATATAAANAATAGAAA	2081
1282	CAGCTTGGAAAAATGGACATTTCCAAGTATGTTTTGGTGGAGTACACAGATATAAANAATAGAAA	1341
2082	TTCTGATGAGAGGTTTCAGTTTTTATATACCAAGTCTCTTAGAGTCTTTAAACATTTGGCCAGC	2141
1342	TTCTGATGAGAGGTTTCAGTTTTTATATACCAAGTCTCTTAGAGTCTTTAAACATTTGGCCAGC	1401
2142	ATCTGTTTATCAAAATGACATAAATACGTAAACCTATAAGAAATTAAGTTTATTAAT	2196
1402	ATCTGTTTATCAAAATGACATAAATACGTAAACCTATAAGAAATTAAGTTTATTAAT	1456

RESULT 6

US-09-905-999-26

; Sequence 26, Application US/09905999

; Patent No. 6797513

; GENERAL INFORMATION:

; APPLICANT: ULLRICH, Axel

; APPLICANT: NAYLER, Oliver

; TITLE OF INVENTION: CLK PROTEIN KINASES AND RELATED PRODUCTS AND METHODS

; FILE REFERENCE: 038602/0431

; CURRENT APPLICATION NUMBER: US/09/905,999

; CURRENT FILING DATE: 2001-07-17

; PRIOR APPLICATION NUMBER: 09/127,248

; PRIOR FILING DATE: 1999-07-31

; PRIOR APPLICATION NUMBER: PCT/IB97/00946

; PRIOR FILING DATE: 1997-06-17

; PRIOR APPLICATION NUMBER: US 08/877,150

; PRIOR FILING DATE: 1997-06-17

; PRIOR APPLICATION NUMBER: US 60/034,286

```

; PRIOR FILING DATE: 1996-12-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 1549
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-905-999-26

Query Match      46.6%; Score 1097.4; DB 4; Length 1549;
Best Local Similarity 90.6%; Pred. No. 9.9e-266;
Matches 1170; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 85 GTCATTATTTAGAACGAGTCCCTTGAATGAGCGAGATATTCGGGACCGGAGATACGTTG 144
DB 239 GTCACATATTTAGAACGAGATGCTTTGAATGAGAGATATTCGGGACCGGAGATACATG 298
QY 145 ACGAATACAGGAATGACTACTGCGAAGGATATGTTCCCTAGACATATATCACAGAGACATG 204
DB 299 ATGAATACAGAAATGACTACTGCGAAGGATATGTTCCAAAGACATTTACCATAGAGACGTTG 358
QY 205 AAAGGGGTATCGAATCCACTGCGAGTAAATCTTCACTCCGAGCAGGAGAGAGAGTCCCTA 264
DB 359 AAAGCACTTTACCGGATCCATTTGCGAGTAAATCCTCAGTCAGGAGCAGGAGAGACGCCCTA 418
QY 265 AAAGGAGCGCAATAGACACTGTTCAAGTCATCATGTCAGTTCGAAAGAGCCACCGAAGGA 324
DB 419 AGAAGAACGTAATAGACCCCTGTGCGAGTCACTAGTCAGTCGATTCGAGAGCCACCGAAGGA 478
QY 325 AAAGTCCAGGAGTATAGAGGATGATGAGGAGGTCACTGATCTGTCAAAAGTGGAGAG 384
DB 479 AAAGTCCAGGAGTATAGAGGATGATGAGGAGGTCACTGATCTGTCAAAAGTGGAGAG 538
QY 385 TTCTAAGAGCAAGATGAAATCGTGGACACATTTGGGTGAAGGAGCCCTTTGGCAAGTTG 444
DB 539 TTCTAAGAGCAAGATGAAATCGTGGACACATTTAGGTGAAGGAGCCCTTTGGCAAGTTG 598
QY 445 TAGAGTGATTCATGCGCATGAGTGGCATGATGAGTGAAGTGAAGTGAAGTGAAGTGAAG 504
DB 599 TAGAGTGATTCATGCGCATGAGTGGCATGATGAGTGAAGTGAAGTGAAGTGAAGTGAAG 658
QY 505 TAGGCGGTTACCGTGAAGCAGTCGTTTCAGAAATCCAAAGTATTAGAGCATTAAATAGTA 564
DB 659 TAGGACGTTACCGGAGGAGCAGTCGTTTCGAAATCCAAAGTATTGAGCATTGAACAGCA 718
QY 565 CTGATCCCAATAGTGTCTTCGAGTGTCCAGATGCTAGATGTTTGAATCATCATGCTC 624
DB 719 CTGACCCCAACAGTGTCTTCGAGTGTCCAGATGCTAGATGTTTGAATCATCATGCTC 778
QY 625 ATGTTTGTATTGTTTGAACCTACTCGGACCTTAGTACTTACGATTTTCAATTAAGAAACA 684
DB 779 ATGTTTGTATTGTTTGAACCTACTCGGACCTTAGTACTTACGATTTTCAATTAAGAAACA 838
QY 685 GCTTTCTGCCATTTTCAAAATTTGACCAATCAGGAGATGCGGTATCAGATCTGCCAGTCAA 744
DB 839 GTTTTCTGCCATTTTCAAAATTTGACCAATCAGGAGATGCGGTATCAGATCTGCCAGTCTA 898
QY 745 TAAATTTTTCATCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTGT 804
DB 899 TAAATTTTTCATCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTAT 958
QY 805 TTGTGAAGTCTGACTATGTAGTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 864
DB 959 TTGTGAAGTCTGACTATGTAGTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1018
QY 865 TGAATAAACACAGATATCAAAAGTTGTGACTTTTGGAAAGTGCACCGTATGATGATGATGAT 924
DB 1019 TGAATAAACACAGATATCAAAAGTTGTGACTTTTGGAAAGTGCACCATATGACGACGAAATC 1078
QY 925 ACAGTACTTTTGTGTTTACCGGCTACAGAGCTCCCGAGGTCTATTTTGGCTTTAGGTT 984
DB 1079 ATAGTACTTTTGTGTTTACCGGCTACAGAGCTCCCGAGGTCTATTTTGGCTTTAGGTT 1138

```

```

QY 985 GGTCTCAGCCCTTGTGATGTTTGGAGCATAGGTTCGATTTCTTATTGAATATTACCTTGGTT 1044
DB 1139 GGTCTCAGCCCTTGTGATGTTTGGAGCATAGGTTCGATTTCTTATTGAGTACTACCTTGGT 1198
QY 1045 TCACAGTCTTTTCAGACTCATGATAGTAAGAGACACCTGGCAATGATGGAACGATATTAG 1104
DB 1199 TCACAGTCTTTTCAGACCCACGATAGTAAGAGACACCTGGCAATGATGGAACGATATTAG 1258
QY 1105 GACCCATACCAACACATGATTTCAAGAAAACAGAAAACGCAAGTATTTTCCACATAACC 1164
DB 1259 GACCCATACCCACACATATGATCCAGAAAGCAAGAAACGCAAGTATTTTCCACATAACC 1318
QY 1165 AGCTAGATTGGGATGAACACAGTTTCTGCTGGTAGATATGTTAGGAGACGCTGCACAAACCGT 1224
DB 1319 AGCTAGATTGGGACGAGCATAGTTTCACTGGGAGATATGTTAGGAGACGCTGCACAAACCGT 1378
QY 1225 TGAAGGAATTTATGCTTTTGCATGATGAAGAAACATGAGAAACCTGTTTGAACCTGGTTCGAA 1284
DB 1379 TAAAGGAATTTATGCTTTTGCATGATGAAGAAACATGAGAAACCTGTTTGAACCTGGTTCGAA 1438
QY 1285 GAATGTTTAGAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCAATTCGAGCATCCTT 1344
DB 1439 GAATGTTTGGAGTATGACCCAGCGAAGGATCACCTTGGATGAAGCAATTCGAGCATCCTT 1498
QY 1345 TCCTTGACTTTATTAAAAAAGAAATGAATCG 1375
DB 1499 TCCTTGACTTTATTAAAAAAGGAATGAGTGGG 1529

```

```

RESULT 7
US-09-810-671-3
; Sequence 3, Application US/09810671
; Patent No. 6455291
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758
; CURRENT APPLICATION NUMBER: US/09/810,671
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 21234
; TYPE: DNA
; ORGANISM: Human
US-09-810-671-3

```

```

Query Match      46.6%; Score 1097; DB 3; Length 21234;
Best Local Similarity 100.0%; Pred. No. 4.4e-265;
Matches 1097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1228 AGGAATTTATGCTTTTGTGATGAAGAAACATGAGAAACCTGTTTGAACCTGGTTCGAGAA 1287
DB 18138 AGGAATTTATGCTTTTGTGATGAAGAAACATGAGAAACCTGTTTGAACCTGGTTCGAGAA 18197
QY 1288 TGTTAGAATATGATCCCAACTCAAGAAATTTACCTTGGATGAAGCAATTCGAGCATCCTTTCT 1347
DB 18198 TGTTAGAATATGATCCCAACTCAAGAAATTTACCTTGGATGAAGCAATTCGAGCATCCTTTCT 18257
QY 1348 TTGACTTTTAAAAAAGAAATGAATGGGAATCAAGTGGTCTTACTATATATCTTCTCTAGA 1407
DB 18258 TTGACTTTTAAAAAAGAAATGAATGGGAATCAAGTGGTCTTACTATATATCTTCTCTAGA 18317
QY 1408 AGAGATTACTTTAGACGTGTGTCACTCACTAACAATTTCTAATATTTTGTAAACATTTAA 1467
DB 18318 AGAGATTACTTTAGACGTGTGTCACTCACTAACAATTTCTAATATTTTGTAAACATTTAA 18377
QY 1468 TTATTTTGTACAGTAAAGTGTAAATATTTGTATGTTTGTATCAATAGCAATAATTAACCTG 1527
DB 18378 TTATTTTGTACAGTAAAGTGTAAATATTTGTATGTTTGTATCAATAGCAATAATTAACCTG 18437

```


Db 19038 TAACATGGCCAGCATCTGTTTATCAAAATGACATAAAATACGTAAACCTTATAAGAAATTAAG 19097
Qy 2188 TTTATTAAATAGGCAATTTATGCTGATTAATCTTACGGGAGAAAGAGGATTTGATTG 2247
Db 19098 TTTATTAAATAGGCAATTTATGCTGATTAATCTTACGGGAGAAAGAGGATTTGATTG 19157
Qy 2248 GAAAGCAGTTTGGGAGAAAGTGCTGCTGCAAAATTTCCAGAAATTTAAATTTGATTGGTTACAT 2307
Db 19158 GAAAGCAGTTTGGGAGAAAGTGCTGCTGCAAAATTTCCAGAAATTTAAATTTGATTGGTTACAT 19217
Qy 2308 AAACCTTTTGACTTCAG 2324
Db 19218 AAACCTTTTGACTTCAG 19234
RESULT 9
US-10-339-656-3
; Sequence 3, Application US/10339656
; Patent No. 6733978
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758D1V2
; CURRENT APPLICATION NUMBER: US/10/339,656
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/109,854
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 21234
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-339-656-3
Query Match 46.6%; Score 1097; DB 4; Length 21234;
Best Local Similarity 100.0%; Pred. No. 4.4e-265;
Matches 1097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1228 AGGAATTTATGCTTTGTCATGATGAAGAACATGAGAACTGTTTGACCTGGTTGGAAGAA 1287
Db 18138 AGGAATTTATGCTTTGTCATGATGAAGAACATGAGAACTGTTTGACCTGGTTGGAAGAA 18197
Qy 1288 TGTTAGAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTCGACGATCCTTTCT 1347
Db 18198 TGTTAGAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTCGACGATCCTTTCT 18257
Qy 1348 TTGACTTATTAAAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATATCTTCTAGA 1407
Db 18258 TTGACTTATTAAAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATATCTTCTAGA 18317
Qy 1408 AGAGATTAATTAAGACTGTGTCAGTCAACTAAACATTCCTAATATATTTTGTAAACATTA 1467
Db 18318 AGAGATTAATTAAGACTGTGTCAGTCAACTAAACATTCCTAATATATTTTGTAAACATTA 18377
Qy 1468 TTATTTTGTACAGTTAAGTGAATATGTTGTTTGTATCAATAGCATAATTAACATTG 1527
Db 18378 TTATTTTGTACAGTTAAGTGAATATGTTGTTTGTATCAATAGCATAATTAACATTG 18437
Qy 1528 TTAAGCAATGATGGTCTTGATAATGCATTAGAAAAATTAATAATTTCTTTTTCGAA 1587
Db 18438 TTAAGCAATGATGGTCTTGATAATGCATTAGAAAAATTAATAATTTCTTTTTCGAA 18497
Qy 1588 ATTACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATTAATGTGATGATC 1647
Db 18498 ATTACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATTAATGTGATGATC 18557

Qy 1648 TTGCTTTTGTACATGGAGGTCACTCTGAAGTGATTTTTTTTGTAGTAAAAAGAAATCTT 1707
Db 18558 TTGCTTTTGTACATGGAGGTCACTCTGAAGTGATTTTTTTTGTAGTAAAAAGAAATCTT 18617
Qy 1708 GACTACTTTATATCTTAAAGGAATATCTTTATATATCTTCAAAATTTAGAACTTAACCTT 1767
Db 18618 GACTACTTTATATCTTAAAGGAATATCTTTATATATCTTCAAAATTTAGAACTTAACCTT 18677
Qy 1768 AAAAGTTTCTTCTGTAATTTGTTGAACGGGTGATTTATTAATCTAGATAAGACAGT 1827
Db 18678 AAAAGTTTCTTCTGTAATTTGTTGAACGGGTGATTTATTAATCTAGATAAGACAGT 18737
Qy 1828 ACTAGAAACCAAAACTCAGAAAAATGTTTACTGTTAGAAATCTTATTAATTTTAAAGTGTG 1887
Db 18738 ACTAGAAACCAAAACTCAGAAAAATGTTTACTGTTAGAAATCTTATTAATTTTAAAGTGTG 18797
Qy 1888 TATTCTTTTTCATTTGGGTGATGTCAGGTGATATACAGACATTCATATGAAAAGGATGAG 1947
Db 18798 TATTCTTTTTCATTTGGGTGATGTCAGGTGATATACAGACATTCATATGAAAAGGATGAG 18857
Qy 1948 TTTCTCCATTTGTGACAGTTTGTAAATAAAACCAACATACACACTTTATTTAAGATTAAA 2007
Db 18858 TTTCTCCATTTGTGACAGTTTGTAAATAAAACCAACATACACACTTTATTTAAGATTAAA 18917
Qy 2008 TCTAACTGGAAAGTCAGCTTGAAAAATGGACATTTCCAAAGTATGTTTGGTGAGTCACAGA 2067
Db 18918 TCTAACTGGAAAGTCAGCTTGAAAAATGGACATTTCCAAAGTATGTTTGGTGAGTCACAGA 18977
Qy 2068 TATAAAATAGAAATTCGTATGAGAGTTTCAGTTTTTAAATACCAAGTCCCTTAGAGTCT 2127
Db 18978 TATAAAATAGAAATTCGTATGAGAGTTTCAGTTTTTAAATACCAAGTCCCTTAGAGTCT 19037
Qy 2128 TAACATTTGGCCAGCATCTGTTTATCAATGACATATAATACCTATAAGAAATTAAG 2187
Db 19038 TAACATTTGGCCAGCATCTGTTTATCAATGACATATAATACCTATAAGAAATTAAG 19097
Qy 2188 TTTATTAATTAGGCAATTTATGCTGTGATTAATCTTACGGGAGAAAGAGGATTTGATTG 2247
Db 19098 TTTATTAATTAGGCAATTTATGCTGTGATTAATCTTACGGGAGAAAGAGGATTTGATTG 19157
Qy 2248 GAAAGCAGTTTGGGAGAAAGTGTGCTGTAATAATTTCCAGAAATTTAAATTTGATTGGTACAT 2307
Db 19158 GAAAGCAGTTTGGGAGAAAGTGTGCTGTAATAATTTCCAGAAATTTAAATTTGATTGGTACAT 19217
Qy 2308 AAACCTTTTGACTTCAG 2324
Db 19218 AAACCTTTTGACTTCAG 19234
RESULT 10
US-09-919-039-238
; Sequence 238, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 238
; LENGTH: 2254
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 420527.25
; NAME/KEY: unsure

Db 623 ACAAGGAAGCGCCGACTAGAAATCAACGCTGCGGAGAAATCAATGAGAAAGATCCTG 682
Qy 574 ATAGTGTCTTCGATGTGTCCAGATGCTAGATGTTGATCATCATGCTGATGTTGTA 633
Db 683 ACAACAGAACTCTGTGTCCAGATGTTGATGTTGATCTACCATGGCCACATGTGTA 742
Qy 634 TTGTTGTTGAATCTGCGGACTTAGTACTTACGATTTCAATTAAGAAAGAAAGCTTTCG 693
Db 743 TCTCTTTGAGCTTCTGGGCTTAGCCTTCCGATTTCTCAAGAGCAACAACTACCTGC 802
Qy 694 CATTTCAATTAACACATCAGGAGATGGGTATCAGATCTGCCAGTCAATTAATTTT 753
Db 803 CCTACCCCATCCCAAGTGGCCATATGCCCTTCCAGCTTCCAGGCGCTGCAAGTTCC 862
Qy 754 TACATCAATAAATAACCCATACAGATCTGAAGCTTGAAGAAATATTTTGTGTAAGT 813
Db 863 TCCATGATAACAGTTGACATACGGACCTCAACCTGAAATATCTGTTTGTGAAT 922
Qy 814 CTGACTATGTAGTCAATATATTAATCTAAATGAACGCTGATGAACGACATGAAACA 873
Db 923 CAGACTACGAACTACCTACAACTTAGAGAAAGAGAGATGAGCGCAGTGTAAAGACA 982
Qy 874 CAGATATCAAGATTTGACATTTGAACTGCAACGTAATGATGATGAACATCACAGTACT 933
Db 983 CAGCGTGGGGTGGTGGACCTTGGCAGTGGCCATTTGACCAAGAACACCATAGCACCA 1042
Qy 934 TGGTGTCTACCGGCACTACAGAGCTCCCGAGGTCAATTTGGCTTTAGTGGTCTCAGC 993
Db 1043 TTGTCTCACTCGCAATTAACGAGCCCGAGGTCACTTGGAGTTGGGCTGGTCCACAGC 1102
Qy 994 CTGTGTGTTGGAGCATAGGTTGCAATCTTAATGAATATTAATCTGTTGTTCAAGTCT 1053
Db 1103 CATGCGATGTAGGAGCATAGGCTGCATCATCTTTGAGTACTACGTTGGCTTCAACCTCT 1162
Qy 1054 TTCAGACTCATGATAGTAAGAGACCTGCAATGATGGAACGATATAGGACCCATAC 1113
Db 1163 TCCAGACCCATGACAAAGAGAGACTTAGCCATGATGGAAGGATCTCGGCTCTGTC 1222
Qy 1114 CACAACATGATTCAGAAAAACAGAAACGCAAGTATTTTCAACATAACCAAGCTAGATT 1173
Db 1223 CTCTCGATGATCAGAAAGACAGAAACAGAAATATTTTATCGGGCTCGCTGGATT 1282
Qy 1174 GGGATGAACACAGTTCTCTGTTGATATGTTAGGAGACGCTGAAACCGTTGAAGGAAT 1233
Db 1283 GGGATGAACACCTCAGCGCGCGCTACGTTCTGTGAGAACTGCAAACTCTGCGGCGGT 1342
Qy 1234 TTATGCTTTGTCATGATCAAGACATGAGAACTGTTTGAACCTGTTTCCGAAAGATGTAG 1293
Db 1343 ATCTGACCTCAGGAGCAGGACACCAACCACTCTCGATCTGATGAAATATGCTAG 1402
Qy 1294 AATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCATTCGAGCATCTTTTCTTGACT 1353
Db 1403 AGTATGAGCTGCTAAGCGGCTGACCTTAGTGAAGCCCTTCAGCATCTTTCTTGCGCT 1462

RESULT 12

US-09-905-999-24
; Sequence 24, Application US/09905999
; Patent No. 6797513
; GENERAL INFORMATION:
; APPLICANT: ULLRICH, Axel
; APPLICANT: NAYLER, Oliver
; TITLE OF INVENTION: CLK PROTEIN KINASES AND RELATED PRODUCTS AND METHODS
; FILE REFERENCE: 038602/0431
; CURRENT APPLICATION NUMBER: US/09/905,999
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 09/127,248
; PRIOR FILING DATE: 1999-07-31
; PRIOR APPLICATION NUMBER: PCT/IB97/00946
; PRIOR FILING DATE: 1997-06-17
; PRIOR APPLICATION NUMBER: US 08/877,150
; PRIOR FILING DATE: 1997-06-17
; PRIOR APPLICATION NUMBER: US 60/034,286

; PRIOR FILING DATE: 1996-12-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
; LENGTH: 1787
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-905-999-24

Query Match 19.3%; Score 454.8; DB 4; Length 1787;
Best Local Similarity 65.0%; Pred. No. 3.7e-104; Indels 0; Gaps 0;
Matches 672; Conservative 0; Mismatches 362

Qy 317 CCGAAGGAAAAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGTGATCTGTCAAAG 376
Db 447 CAGTAAGCGCAGCAGCCGAGTGTGGAAGATGACAAGGAGGCCACCTGGTGTGCCGAT 506
Qy 377 TGGAGACGTTCTAAGACAGATATGAAATCGTGTGAACCTTTGGTGAAGAGCCTTTGG 436
Db 507 CGGCGATTTGGCTCCAAGAGCGATATGATGATCGTGGGAAACCTGGGTGAAGCACCCTTTGG 566
Qy 437 CAAAAGTTGTAGAGTCATTGATCATGTCATGCGATGCGATGATGATGATGATGATGATGAT 496
Db 567 CAAAGTGTGAGTGCTTGGACCATGCCAGAGGGAAGTCAACAGTTGGCTTGAAGATCAT 626
Qy 497 AAAAAATGTAGCGCTTACCGTGAAGCAGCTCGTTCAAAAAATCAAGATTTAGAGCATT 556
Db 627 CCGTAATGTGGGCAAGTATCGGAAGCTCGTCTGTAGAAATTAATGTCTCAAGAAAT 686
Qy 557 AAATAGTACTGATCCCAATAGTGTCTCCGATGTGTCAGATGTCAGATGTCAGATGTCAGTCA 616
Db 687 CAAAGGAAAGACAAGGAAATAAGTTCTTGTCTGATGTCGACTGGTTCACCTT 746
Qy 617 TCATGTCATCTTTGTTGTTGTTGAACTACTCGGACTTAGTACTTACGATTTTCAATTA 676
Db 747 CCATGTCATATGTCATCGCTTTGAGCTCTGGGCAAGAACACCTTTGATGTTCTGAA 806
Qy 677 AGAAAAACAGCTTCTGCGCATTTCAAATTTGACCAATCAGGAGATGCGGTATCAGATCTG 736
Db 807 GGAGAACAACTTCCAGCCTTACCCTTACCACATGTCGGGCACATGGCTTACCAGCTCTG 866
Qy 737 CCAGTCAATAAATTTTACATCATATAAATTAACCATACAGATCTGAAGCTTGAAGA 796
Db 867 TCATGCGCTTGAATTTCTACACGAGAACCAAGCTACCCACACAGACTTGAAGCCAGAGAA 926
Qy 797 TATTTTGTGTTGTAAGTCTGACTATGATCAAAATAAATTTCTAAAAATGAAACGCTGTA 856
Db 927 CATCTTGTGTTGTAATTTCTGATTTGAAACCTCTCTA CAATGAGCACAGAGCTGCGAGGA 986
Qy 857 ACGCACACTGAAAAACACAGATATCAAAGTTGTTGACTTTTGGAAAGTGCACAGTATGATGA 916
Db 987 GAAGTCAGTGAAGAAACCAAGCATCCGAGTGGCAGACTTTGGCAGTGCACAGTTTGACCA 1046
Qy 917 TGAACATCAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 976
Db 1047 TGAACATCAGACCACTTGGCCACCCCTGCTACCGCCACCTGAGGTGATCCTTGA 1106
Qy 977 TTTAGGTTGGTCTCAGCGCTTGTGATGTTGGAGCATAGTTGCTTCTTATTGAATATA 1036
Db 1107 GCTGGGCTGGGCACAGCTTGTGATGCTGGAGTATCGGCTGCATTTCTTTTGATGACTA 1166
Qy 1037 CTTTGGTTTCAAGTCTTTCAAGTCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1096
Db 1167 CCGTGGCTTTACACTCTTCCAGACCCATGAAATAGAGAACACTTGGTTATGATGAGGAA 1226
Qy 1097 AATATTAGGCCCCATACCAACACATGATTTCAAGAAACAGAAACCGCAAGTATTTTCA 1156
Db 1227 GATTTAGGACCCATCCCATGATCCACATGATCCACCGTACCAAGGAGCAAGAAATATTTCTA 1286
Qy 1157 CCATAACCAAGCTAGATTTGGGATGAACAGTCTTCTGCTGGTAGATATGTTAGGAGACGCTG 1216
Db 1287 CAAAGGGGGCTGTTTGGGATGAGAACAGCTCTGATGGGCGGTATGTTGAAGGAGAACTG 1346


```

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2648
; LENGTH: 1763
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-2648

Query Match      18.3%; Score 430.8; DB 4; Length 1763;
Best Local Similarity 63.5%; Pred. No. 3.9e-98;
Matches 657; Conservative 0; Mismatches 377; Indels 0; Gaps 0;

Qy 317 CCGAAGGAAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAG 376
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 440 CAGTAAGCGCAGCAGCCGGAGTGTGGAAGATGACAAGGAGGTCACCTGTGTGCGGAT 499
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 377 TGGAGACGTTCTAAGAGCAAGATATGAATCGTGACACTTTGGTGAAGGACCTTTGG 436
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 500 CGGCGATTGGCTTCCAGAGCGATATGAGTTGTGGGAACTGGGTGAAGCACCCTTTGG 559
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 437 CAAAGTGTAGAGTCATTTAGTCATGCGATGCGCATGCGATGCGATGCGATGCGATGCG 496
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 560 CAAGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 619
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 497 AAAAAATGTAGCGCTTACCGTGAAGCAGCTCGTTTCAAGAAATCAAGTATTAGAGCACTT 556
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 620 CCGCAACGTGGGCAAGTACCGGGAGGTCGCCGGCTAGAAATCAACGTGCTCAAAAAAT 679
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 557 AAATAGTACTATCCCAATAGTGTCTCCGATGTGTCCAGATGCTAGATGCTGTGATGCT 616
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 617 TCATGGTTCATGTTTGTATGTTGTAATACTAGTACTAGTACTAGTACTAGTACTAGTACT 676
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 740 CCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 799
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 677 AGAAACAGCTTTTGGCAATTTCAAATTTGACCACTACCGGACATGTCGGGACATGTCG 736
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 800 GGAGAAATAACTTCCAGGCTTACCCCTTACCAATGTCGGGACATGTCGGGACATGTCG 859
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 737 CCAGTCAATAAATTTTACATCAATAAATTAACCCATACAGATCTGAAGCCTGAAAA 796
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 860 CCAAGGAGGAGCAAGAAACAAGTTCCTGATGAGATCAGTCAACCATACAGATCTGA 919
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 797 TATTTTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 856
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 920 CATCTGTTGTGAATCTGAGTTTGAACCTCTACAAATGAGCAGCAGAGCTGTGAGGA 979
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 857 ACGCACACTGAAACACAGATATCAAAAGTTTGTGACCTTTGGAAGTGAACGTATATGA 916
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 980 GAAGTCAGTGAAGAACACACAGATCCGAGTGGCTGACTTTTGGCAGTGCACATTTGACCA 1039
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 917 TGAACATCACAGTACTTTGGTGTCTACCCGCACTTACAGAGCTCCCGAGGTCATTTTGGC 976
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1040 TGAGCACACACACCAATTTGGCCACCCCTCTATCTATCCCGCTGAGGTGATCCCTTGA 1099
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 977 TTTAGGTTGGTCTCAGCCTCTGTGATGTTTGGAGCATAGGTTGCAATCTTATTAATAATA 1036
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1100 GCTGGGCTGGGCACAGCCCTGTGAGCTGTGAGCATTTGGCTGCAATCTCTTTTGTAGTACTA 1159
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1037 CTTTGGTTTACAGTCTTTTACAGTCTATGATAGTAAGAGCAGCCTGGCAATGGAACG 1096
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1160 CCGGGGCTTCACTCTTCCAGACCCACGAAAA CCGAGAGCAGCCTGGTGTGATGGAGAA 1219
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Qy 1097 AATATTAGGACCCATACACAAACATGATTACAGAAACAAAGAAACGCAAGTATTTTCA 1156
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1220 GATCTAGGCGCCATCCATCACAATGATCCACCGTACAGGAGCAGAAATATTTCTA 1279
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1157 CCATAACGAGTATGATTTGGGATGAACACAGTTCTGCTGGTATGATGTTAGAGACGCTG 1216
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1280 CAAAGGGGCGCTAGTTTGGGATGAGAACAGCTCTGACGCGCGGTATGTGAAGAGAACTG 1339
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1217 CAAACCGTTGAAGGAATTTATGCTTTTGTATGATGAAGAACATGAGAACTGTTTGACCT 1276
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1340 CAAACCTCTGAAGAGTTACATGCTCCCAAGATCTCCCTGGAGCAGCTGTTTGACCT 1399
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1277 GGTTCGAAGATTTAGAAATATGATCCAACTCAAGAAATTTACCTTTGATGAAGCAATTGCA 1336
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1400 GATGAGGAGGATTTAGAAATTTGACCTGCCAGGCACTACACTGGCCGAGGCGCTGCT 1459
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1337 GCATCCTTTCTTTG 1350
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1460 GCACCCCTTCTTTG 1473
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
US-09-949-016-2649
; Sequence 2649, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2649
; LENGTH: 1763
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-2649

Query Match      18.3%; Score 430.8; DB 4; Length 1763;
Best Local Similarity 63.5%; Pred. No. 3.9e-98;
Matches 657; Conservative 0; Mismatches 377; Indels 0; Gaps 0;

Qy 317 CCGAAGGAAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAG 376
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 440 CAGTAAGCGCAGCAGCCGGAGTGTGGAAGATGACAAGGAGGTCACCTGTGTGCGGAT 499
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 377 TGGAGACGTTCTAAGAGCAAGATATGAATCGTGACACTTTGGTGAAGGACCTTTGG 436
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 500 CGGCGATTGGCTTCCAGAGCGATATGAGTTGTGGGAACTGGGTGAAGCACCCTTTGG 559
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 437 CAAAGTGTAGAGTCATTTAGTCATGCGATGCGCATGCGATGCGATGCGATGCGATGCG 496
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 560 CAAGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 619
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 497 AAAAAATGTAGCGCTTACCGTGAAGCAGCTCGTTTCAAGAAATCAAGTATTAGAGCACTT 556
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 620 CCGCAACGTGGGCAAGTACCGGGAGGTCGCCGGCTAGAAATCAACGTGCTCAAAAAAT 679
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 557 AAATAGTACTATCCCAATAGTGTCTCCGATGTGTCCAGATGCTAGATGCTGTGATGCT 616
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 617 TCATGGTTCATGTTTGTATGTTGTAATACTAGTACTAGTACTAGTACTAGTACTAGTACT 676
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 740 CCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 799
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 677 AGAAACAGCTTTTGGCAATTTCAAATTTGACCACTACCGGACATGTCGGGACATGTCG 736
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 800 GGAGAAATAACTTCCAGGCTTACCCCTTACCAATGTCGGGACATGTCGGGACATGTCG 859
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 737 CCAGTCAATAAATTTTACATCAATAAATTAACCCATACAGATCTGAAGCCTGAAAA 796
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 860 CCAAGGAGGAGCAAGAAACAAGTTCCTGATGAGATCAGTCAACCATACAGATCTGA 919
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 797 TATTTTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 856
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 920 CATCTGTTGTGAATCTGAGTTTGAACCTCTACAAATGAGCAGCAGAGCTGTGAGGA 979
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 857 ACGCACACTGAAACACAGATATCAAAAGTTTGTGACCTTTGGAAGTGAACGTATATGA 916
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 980 GAAGTCAGTGAAGAACACACAGATCCGAGTGGCTGACTTTTGGCAGTGCACATTTGACCA 1039
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 917 TGAACATCACAGTACTTTGGTGTCTACCCGCACTTACAGAGCTCCCGAGGTCATTTTGGC 976
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1040 TGAGCACACACACCAATTTGGCCACCCCTCTATCTATCCCGCTGAGGTGATCCCTTGA 1099
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 977 TTTAGGTTGGTCTCAGCCTCTGTGATGTTTGGAGCATAGGTTGCAATCTTATTAATAATA 1036
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1100 GCTGGGCTGGGCACAGCCCTGTGAGCTGTGAGCATTTGGCTGCAATCTCTTTTGTAGTACTA 1159
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1037 CTTTGGTTTACAGTCTTTTACAGTCTATGATAGTAAGAGCAGCCTGGCAATGGAACG 1096
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1160 CCGGGGCTTCACTCTTCCAGACCCACGAAAA CCGAGAGCAGCCTGGTGTGATGGAGAA 1219
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Qy	677	AGAAAAACAGCTTTCTGCCAATTTCAAAATTGACCAATCAGGCAGATGGCGTATCAGATCTG	736
Db	800	GGAGAATAACTTCAGCCTTACCCCTACCAATGTCGGCACATGGCTACCAGCTCTG	859
Qy	737	CCAGTCAATAATTTTACATATAATAAATTAAACCATACAGATCTGAAGCCTGAAAA	796
Db	860	CCAGCCTTATAGATTCTGCAATGAATCAGCTGACCCATACAGACTTGAACCCAGAGAA	919
Qy	797	TATTTTGTGTGAAGTCTGACTATGTAGTCAAAATATAAATTCTAAAAATGAAACGTGATGA	856
Db	920	CATCTGTGTGTGATTTCTGAGTTTGAACCTCTACAAATGAGCACAGAGCTGTGAGGA	979
Qy	857	ACGCACACTGAAAAACACAGATATCAAAAGTTGTGACTTTGGAAGTGCACAGTATGATGA	916
Db	980	GAAGTCAGTGAAGAACACAGCATCCGAGTGGCTGACTTTGGCAGTGCACATTTGACCA	1039
Qy	917	TGAACATCACAGTACTTTGGTGTCTACCCGGCACTACAGAGCTCCCGAGGTCAATTTGGC	976
Db	1040	TGAGCACACACCCCAATTTGGCCACCCCGTCACTATCCGCCCGCTGAGGTGATCCTTGA	1099
Qy	977	TTTAGTTGGTCTCAGCCTGTGTGATGTTTGGAGCATAGTTTGCATTTCTATTGAATATTA	1036
Db	1100	GCTGGGCTGGGCACAGCCCTGTGAGCTCTGAGGCAATGGCTGCATTTCTTTGAGTACTA	1159
Qy	1037	CTTTGGTTTCAAGTCTTTTCAGACTCATGATAGTAAGAGACACCTGGCAATGATGGACG	1096
Db	1160	CGGGGCTTTCACACTCTTCCAGACCCACGAAAAACCGAGAGCACCTGGTGTATGATGGAGAA	1219
Qy	1097	AATATTAGGACCCATACACACACATGATTCAGAAACAGAAAAACGCAAGTATTTTCA	1156
Db	1220	GATCCTAGGGCCCATCCCATCACATGATCCCGTACCGAGGAGCAGAAAAATATTTCTA	1279
Qy	1157	CCATAACCAAGCTAGATTGGGATGAACACAGTTCTGCTGTAGATATGTTAGGAGAGCTG	1216
Db	1280	CAAGGGGGCCCTAGTTTGGGATGAGAACAGCTCTGACGGCCGGTATGTGAAGGAGACTG	1339
Qy	1217	CAAAACCGTTGAAGGAATTTATGCTTTGTGATGATGAAGAACATGAGAAACTGTTTGACCT	1276
Db	1340	CAAACTCTGAAGAGTTTACATGCTCCAAGACTCCCTGGAGCACGTCAGCTGTTTGACCT	1399
Qy	1277	GGTTGGAAGATGTAGAAATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCAATTGCA	1336
Db	1400	GATGAGGAGGATGTAGAAATTTGACCTTGCCCGAGCGCATCACACTGGCCGAGGCCCTGCT	1459
Qy	1337	GCATCCTTTCTTTG	1350
Db	1460	GCACCCCTTCITTG	1473

Search completed: March 16, 2005, 18:53:25
Job time : 400.184 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw.model

Run on: March 16, 2005, 14:32:04 ; Search time 1314.36 Seconds
(without alignments)

10662.144 Million cell updates/sec

Title: US-10-801-671-1

Perfect score: 2354

Sequence: 1 gccagctgggttacttta.....taaaaaaacaaaaaaac 2354

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5544816 seqs, 2976611598 residues

Total number of hits satisfying chosen parameters: 11089632

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09D_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2354	100.0	2354	9	US-09-810-671-1
2	2354	100.0	2354	13	US-10-109-854-1
3	2354	100.0	2354	15	US-10-339-656-1
4	2354	100.0	2354	18	US-10-801-671-1
5	1908.4	81.1	4035	17	US-10-425-114-26852
6	1286	54.6	1446	17	US-10-267-502-135
7	1178.2	50.1	1456	17	US-10-641-643-699
8	1097.4	46.6	1549	9	US-09-905-999-26
9	1097.4	46.6	1549	18	US-10-825-177-26
10	1097.2	46.6	1446	17	US-10-267-502-138
11	1097	46.6	21234	9	US-09-810-671-3
					Sequence 1, Appli
					Sequence 1, Appli
					Sequence 1, Appli
					Sequence 1, Appli
					Sequence 135, App
					Sequence 699, App
					Sequence 26, Appl
					Sequence 26, Appl
					Sequence 138, App
					Sequence 3, Appli

12	1097	46.6	21234	13	US-10-109-854-3	Sequence 3, Appli
13	1097	46.6	21234	15	US-10-339-656-3	Sequence 3, Appli
14	1097	46.6	21234	18	US-10-801-671-3	Sequence 3, Appli
15	977.4	41.5	3040	17	US-10-425-114-26212	Sequence 26212, A
16	813.2	34.5	1455	17	US-10-267-502-136	Sequence 136, App
17	811.6	34.5	1834	17	US-10-154-708-3	Sequence 3, Appli
18	811.6	34.5	1834	18	US-10-755-889-1	Sequence 1, Appli
19	768.8	32.7	1452	17	US-10-267-502-139	Sequence 139, App
20	768.4	32.6	906	17	US-10-267-502-140	Sequence 140, App
21	644.4	27.4	1743	17	US-10-154-708-10	Sequence 10, Appl
22	640.6	27.2	2254	10	US-09-919-039-238	Sequence 238, App
23	640	27.2	2516	9	US-09-925-298-121	Sequence 121, App
24	640	27.2	2516	14	US-10-102-806-121	Sequence 121, App
25	499.6	21.2	1500	10	US-09-790-852-2	Sequence 2, Appli
26	499.6	21.2	1973	9	US-09-962-436-266	Sequence 266, App
27	499.6	21.2	1973	9	US-09-880-107-2190	Sequence 2190, Ap
28	499.6	21.2	1973	18	US-10-737-450-131	Sequence 131, App
29	481.2	20.4	1538	9	US-09-905-999-22	Sequence 22, Appl
30	481.2	20.4	1538	18	US-10-825-177-22	Sequence 22, Appl
31	464.4	19.7	1296	17	US-10-182-243-16	Sequence 16, Appl
32	464.4	19.7	1628	17	US-10-425-114-26266	Sequence 26266, A
33	461.6	19.6	1026	17	US-10-267-502-133	Sequence 133, App
34	458.6	19.5	1885	17	US-10-439-703-89	Sequence 89, Appl
35	454.8	19.3	1787	9	US-09-905-999-24	Sequence 24, Appl
36	454.8	19.3	1787	18	US-10-825-177-24	Sequence 24, Appl
37	453.2	19.3	1473	17	US-10-267-502-137	Sequence 137, App
38	432.4	18.4	1473	17	US-10-267-502-134	Sequence 134, App
39	432.4	18.4	1762	17	US-10-305-720-1439	Sequence 1439, Ap
40	430.8	18.3	2505	17	US-10-108-260A-2256	Sequence 2256, Ap
41	429.2	18.2	2621	17	US-10-104-047-656	Sequence 656, App
42	377.4	16.0	1181	17	US-10-425-114-16170	Sequence 16170, A
43	369	15.7	369	9	US-09-796-692-5456	Sequence 5456, Ap
44	369	15.7	369	14	US-10-040-862-5456	Sequence 5456, Ap
45	369	15.7	369	17	US-10-057-475B-5456	Sequence 5456, Ap

ALIGNMENTS

RESULT 1
US-09-810-671-1
; Sequence 1, Application US/09810671
; Publication No. US20020076783A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO00758
; CURRENT APPLICATION NUMBER: US/09/810.671
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Human
US-09-810-671-1

Query Match	100.0%	Score	2354	DB	9	Length	2354
Best Local Similarity	100.0%	Pred. No.	0				
Matches	2354	Conservative	0	Mismatches	0	Indels	0
Qy	1	GCAGCTGGGTACTTTAA	AAACATGCTCATGTGATCCCTCTTGAAGCTTCGCACT	60			
Db	1	GCAGCTGGGTACTTTAA	AAACATGCTCATGTGATCCCTCTTGAAGCTTCGCACT	60			
Qy	61	CTGTTGAAGAGGACACT	CATCCAGTCAATTTAGAACGAGGTCCTTGAATGAGCGAG	120			
Db	61	CTGTTGAAGAGGACACT	CATCCAGTCAATTTAGAACGAGGTCCTTGAATGAGCGAG	120			
Qy	121	ATTATCGGACCGGAGAT	CGTTGACGAATCAGGAATGACTTGTGAGGATATGTC	180			

QY 2341 AACAAAAAACAAC 2354
DB 2341 AACAAAAAACAAC 2354

RESULT 2

US-10-109-854-1
; Sequence 1, Application US/10109854
; Publication No. US20020119548A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE OF INVENTION: THEROF
; FILE REFERENCE: CLO00758DIV
; CURRENT APPLICATION NUMBER: US/10109,854
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PaetSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-109-854-1

Query Match 100.0%; Score 2354; DB 13; Length 2354;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCAGCTGGGTACTTTAAAAAACATGCTCCATGTCATCCCTCTCGAAGCTTCGCACT 60
DB 1 GCCAGCTGGGTACTTTAAAAAACATGCTCCATGTCATCCCTCTCGAAGCTTCGCACT 60
QY 61 CTGTTGAAGAGGACACTCATCCAGTCATTAATTTAGAAGCAAGGTCTTGAATGAGCGAG 120
DB 61 CTGTTGAAGAGGACACTCATCCAGTCATTAATTTAGAAGCAAGGTCTTGAATGAGCGAG 120
QY 121 ATTATCGGACCGGAGATACGTTGACGAATACAGGAATGACTACTGTGAGGATATGTC 180
DB 121 ATTATCGGACCGGAGATACGTTGACGAATACAGGAATGACTACTGTGAGGATATGTC 180
QY 181 CTAGACATTTATCAGAGACATTTGAAGCGGTATCGAATCCACTGCAATTAATCTTCAG 240
DB 181 CTAGACATTTATCAGAGACATTTGAAGCGGTATCGAATCCACTGCAATTAATCTTCAG 240
QY 241 TCCGAGCAGGAGAGCAGTCTTAAAGGAGCGCAATAGACACTGTTCAAGTCATCAGT 300
DB 241 TCCGAGCAGGAGAGCAGTCTTAAAGGAGCGCAATAGACACTGTTCAAGTCATCAGT 300
QY 301 CACGTTTGAAGAGCAGCAGGAGAAAGATCAGAGATATAGAGGATATAGAGGAGGTC 360
DB 301 CACGTTTGAAGAGCAGCAGGAGAAAGATCAGAGATATAGAGGATATAGAGGAGGTC 360
QY 361 ACCTCATCTCTCAAGTGGAGAGCTTCTAAGAGCAAGATATGAATCGTGGACACTTTGG 420
DB 361 ACCTCATCTCTCAAGTGGAGAGCTTCTAAGAGCAAGATATGAATCGTGGACACTTTGG 420
QY 421 GTGAAGGAGCCTTTGGCAAAAGTTGTAGAGTGCAATTCATGCGCATGGATGCGATGTC 480
DB 421 GTGAAGGAGCCTTTGGCAAAAGTTGTAGAGTGCAATTCATGCGCATGGATGCGATGTC 480
QY 481 TAGCAGTGAATATCGTAAAGATAGGCGGTACCGTGAAGCAGCTGTTGAGAAATCC 540
DB 481 TAGCAGTGAATATCGTAAAGATAGGCGGTACCGTGAAGCAGCTGTTGAGAAATCC 540
QY 541 AAGTATTAGACACTTAATAGTACTGATCCCAATAGTGTCTTCCGATGTGTCAGATGC 600
DB 541 AAGTATTAGACACTTAATAGTACTGATCCCAATAGTGTCTTCCGATGTGTCAGATGC 600

QY 601 TAGAATGGTTTGTATCATCATGGTCAATGTTGTATGTTGTGAACTACTGGGACTTAGTA 660
DB 601 TAGAATGGTTTGTATCATCATGGTCAATGTTGTATGTTGTGAACTACTGGGACTTAGTA 660
QY 661 CTTACGATTTTCATTAAGAAAAACAGCTTTTCGCCAATTTCAAAATTCACCAATCAGGAG 720
DB 661 CTTACGATTTTCATTAAGAAAAACAGCTTTTCGCCAATTTCAAAATTCACCAATCAGGAG 720
QY 721 TGGCGTATCAGATCGCCAGTCAATAAATTTTATCATCAATAAATAAATTAACCCATACAG 780
DB 721 TGGCGTATCAGATCGCCAGTCAATAAATTTTATCATCAATAAATAAATTAACCCATACAG 780
QY 781 ATCTGAAGCCTGAAATAATTTTGTGTAAGTCTGACTATGTAATTAATTTCTA 840
DB 781 ATCTGAAGCCTGAAATAATTTTGTGTAAGTCTGACTATGTAATTAATTTCTA 840
QY 841 AAATGAACGCTGATGAACGACACTGAAAAACACAGATATCAAAAGTTGTCACCTTTGAA 900
DB 841 AAATGAACGCTGATGAACGACACTGAAAAACACAGATATCAAAAGTTGTCACCTTTGAA 900
QY 901 GTGCAACGATGATGATGAACATCACAGTACTTTTGGTGTCTACCCGGCACTACAGAGTC 960
DB 901 GTGCAACGATGATGATGAACATCACAGTACTTTTGGTGTCTACCCGGCACTACAGAGTC 960
QY 961 CCGAGGTCATTTTGGCTTTAGGTTGGTCTCAGCCTTGTGATGTTGGAGCATAGGTTGCA 1020
DB 961 CCGAGGTCATTTTGGCTTTAGGTTGGTCTCAGCCTTGTGATGTTGGAGCATAGGTTGCA 1020
QY 1021 TTTCTATTGAATATTACCTTGGTTTACAGTCTTTTACAGTCTGATAGTAAAGAGCACC 1080
DB 1021 TTTCTATTGAATATTACCTTGGTTTACAGTCTTTTACAGTCTGATAGTAAAGAGCACC 1080
QY 1081 TGGCAATGATGGAACGAATATTAGGACCCATACCAACACATGATTCAGAAAAACAAGAA 1140
DB 1081 TGGCAATGATGGAACGAATATTAGGACCCATACCAACACATGATTCAGAAAAACAAGAA 1140
QY 1141 AACGCAAGTATTTTCAACATTAACAGCTAGATGAGGATGAACACAGTCTGCTGAGTAG 1200
DB 1141 AACGCAAGTATTTTCAACATTAACAGCTAGATGAGGATGAACACAGTCTGCTGAGTAG 1200
QY 1201 ATGTTAGGAGACGCTGCAACCGTTGAGGAATTTATGCTTTGTCATCATGAAGAACATG 1260
DB 1201 ATGTTAGGAGACGCTGCAACCGTTGAGGAATTTATGCTTTGTCATCATGAAGAACATG 1260
QY 1261 AGAACTGTTTGTGACCTGCTTGAAGAAATGTTAGAAATATGATCCAACTCAAGAAATACCT 1320
DB 1261 AGAACTGTTTGTGACCTGCTTGAAGAAATGTTAGAAATATGATCCAACTCAAGAAATACCT 1320
QY 1321 TGGATGAAGCATTTGAGCATCCTTTCTTGTGACTTATTAAGAAAGAAATGAAATGGGAATC 1380
DB 1321 TGGATGAAGCATTTGAGCATCCTTTCTTGTGACTTATTAAGAAAGAAATGAAATGGGAATC 1380
QY 1381 AGTGGTCTTACTATATCTTCTAGAGAGATTTAAAGCAATGTTAGAAATGAGTCAACTAAA 1440
DB 1381 AGTGGTCTTACTATATCTTCTAGAGAGATTTAAAGCAATGTTAGAAATGAGTCAACTAAA 1440
QY 1441 CATTTCTAATAATTTTGTAAACATTAATTTTGTGACAGTTAAATGTAATTTGATG 1500
DB 1441 CATTTCTAATAATTTTGTAAACATTAATTTTGTGACAGTTAAATGTAATTTGATG 1500
QY 1501 TTTTGTATCAATAGCATTAATTAATTTTGTAAAGCAAGTATGCTTTGATTAATGCAATTAGAA 1560
DB 1501 TTTTGTATCAATAGCATTAATTAATTTTGTAAAGCAAGTATGCTTTGATTAATGCAATTAGAA 1560
QY 1561 AAATTAATTAATTTTCTTTTGAATTAACCTTTTAAATACCTTTTGAATATCCCT 1620
DB 1561 AAATTAATTAATTTTCTTTTGAATTAACCTTTTAAATACCTTTTGAATATCCCT 1620
QY 1621 TGTGTCAGTGAATAATGATGATCTTGTGCTTTTGTACATGAGAGTCACTCTGGAAGT 1680
DB 1621 TGTGTCAGTGAATAATGATGATCTTGTGCTTTTGTACATGAGAGTCACTCTGGAAGT 1680
QY 1681 GATTTTGTGAGTAAAGGAAATCTTGCATCTTATATTTCTTAAAGAAATATTTCTTTA 1740


```

1681 GATTTTGTGTAAGGAAATCTTGACTACTTTATATCTTAAAGGAATATCTTTA 1740
1741 TATACCTCAAAATTTAGAACTTAACTTTAAAGTTTCTTCTGTAATCTGCAACGGGTG 1800
1741 TATACCTCAAAATTTAGAACTTAACTTTAAAGTTTCTTCTGTAATCTGCAACGGGTG 1800
1801 ATTATTTAACTCTAGATAAAGCACTAGATAAAGCACTAGATAAAGCACTAGATAAAGCA 1860
1801 ATTATTTAACTCTAGATAAAGCACTAGATAAAGCACTAGATAAAGCACTAGATAAAGCA 1860
1861 TAGAATCTTAAATTTAACTTTAACTTTAACTTTAACTTTAACTTTAACTTTAACTTTAA 1920
1861 TAGAATCTTAAATTTAACTTTAACTTTAACTTTAACTTTAACTTTAACTTTAACTTTAA 1920
1921 ACCAGACATTTCAATGAAAGGCACTGAGTTTGTCCATTTGTGACAGTTTGTGTTTAAATAAAC 1980
1921 ACCAGACATTTCAATGAAAGGCACTGAGTTTGTCCATTTGTGACAGTTTGTGTTTAAATAAAC 1980
1981 ACATACACACTTTTAAAGATTTAAATCTAACTGAAAGTCACTGAAAGTCACTGAAAGTCA 2040
1981 ACATACACACTTTTAAAGATTTAAATCTAACTGAAAGTCACTGAAAGTCACTGAAAGTCA 2040
2041 TTCCAGATGTTTGGTGAAGTCAAGATATAAATAAATAAATAAATAAATAAATAAATAAATA 2100
2041 TTCCAGATGTTTGGTGAAGTCAAGATATAAATAAATAAATAAATAAATAAATAAATAAATA 2100
2101 TTTTAAATACCAAGTCTTAAAGATTTAAATCTAACTGAAAGTCACTGAAAGTCACTGAAAG 2160
2101 TTTTAAATACCAAGTCTTAAAGATTTAAATCTAACTGAAAGTCACTGAAAGTCACTGAAAG 2160
2161 TAAATACGTAACCTTAAAGATTTAAATCTAACTGAAAGTCACTGAAAGTCACTGAAAGTCA 2220
2161 TAAATACGTAACCTTAAAGATTTAAATCTAACTGAAAGTCACTGAAAGTCACTGAAAGTCA 2220
2221 TCTTACGGGAGAAAGAGATTTGATTTGGAAGAGATTTGGAAGAGATTTGGAAGAGATTTG 2280
2221 TCTTACGGGAGAAAGAGATTTGATTTGGAAGAGATTTGGAAGAGATTTGGAAGAGATTTG 2280
2281 TTCCAGAAATTTAAATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 2340
2281 TTCCAGAAATTTAAATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 2340
2341 AACAAAAAATAAAC 2354
2341 AACAAAAAATAAAC 2354

```

```

RESULT 3
US-10-339-656-1
; Sequence 1, Application US/10339656
; Publication No. US20030134319A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758DIV2
; CURRENT APPLICATION NUMBER: US/10/339,656
; PRIORITY FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/109,854
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-339-656-1

```

```

Query Match
Best Local Similarity 100.0%; Score 2354; DB 15; Length 2354;
Matches 2354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCAGCTGGGTTACTTTTAAAAAATCTGCTCAATGTGATCCCTTTGAAGCTTCGCACT 60
DB 1 GCCAGCTGGGTTACTTTTAAAAAATCTGCTCAATGTGATCCCTTTGAAGCTTCGCACT 60
QY 61 CTGTTGAGAGGACACTCATCCAGTCAATTTTAAAGAGGAGGCTTCTTGAATGAGCGAG 120
DB 61 CTGTTGAGAGGACACTCATCCAGTCAATTTTAAAGAGGAGGCTTCTTGAATGAGCGAG 120
QY 121 ATTATCGGCAACCGGAGATACCTTGACGAATACAGGAATGACTACTGTGCAAGGATATGTC 180
DB 121 ATTATCGGCAACCGGAGATACCTTGACGAATACAGGAATGACTACTGTGCAAGGATATGTC 180
QY 181 CTAGACATTTATCACAGAGACATTTGAAAGCGGATATCGAATCCACTGCAAGTAAATCTTCAG 240
DB 181 CTAGACATTTATCACAGAGACATTTGAAAGCGGATATCGAATCCACTGCAAGTAAATCTTCAG 240
QY 241 TCCGAGCAGGAGAGCAGTCTTAAAGAGGAGCGCAATAGACACTGTTTCAAGTCAATCAGT 300
DB 241 TCCGAGCAGGAGAGAGAGTCTTAAAGAGGAGCGCAATAGACACTGTTTCAAGTCAATCAGT 300
QY 301 CACGTTTCAAGAGCCACCGAAGGAAAGATCCAGAGGTATAGAGGATGATGAGGAGGTC 360
DB 301 CACGTTTCAAGAGCCACCGAAGGAAAGATCCAGAGGTATAGAGGATGATGAGGAGGTC 360
QY 361 ACCTGATCTGCTCAAGTGGAGAGCGTTCTTAAGAGCAAGATATGAAATCTGTCGACATTTGG 420
DB 361 ACCTGATCTGCTCAAGTGGAGAGCGTTCTTAAGAGCAAGATATGAAATCTGTCGACATTTGG 420
QY 421 GTGAGGAGCGCTTTGGCAAGTCTGAGAGTCAATGATTCATGCGCATGATGCGCATGATG 480
DB 421 GTGAGGAGCGCTTTGGCAAGTCTGAGAGTCAATGATTCATGCGCATGATGCGCATGATG 480
QY 481 TAGCAGTGAATAATCGTAAAAAATGAGGCGGTTACCGTGAAGAGAGTCTGTTTCAAGAAATCC 540
DB 481 TAGCAGTGAATAATCGTAAAAAATGAGGCGGTTACCGTGAAGAGAGTCTGTTTCAAGAAATCC 540
QY 541 AAGTATTAGAGCCTTAAATAGTACTGATCCCAATAGTGTCTTCCGATGTCAGATGTC 600
DB 541 AAGTATTAGAGCCTTAAATAGTACTGATCCCAATAGTGTCTTCCGATGTCAGATGTC 600
QY 601 TAGAATGTTTGTATCATCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 660
DB 601 TAGAATGTTTGTATCATCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 660
QY 661 CTATACGATTTTCAATTAAGAAACAGCTTTCTGCAATTTCAAAATTTGACCACTACAGGAGA 720
DB 661 CTATACGATTTTCAATTAAGAAACAGCTTTCTGCAATTTCAAAATTTGACCACTACAGGAGA 720
QY 721 TGGCGTATCAGATCTGCGAGTCAATAATTTTACATCATATAAATTAATTAATTAATTAAT 780
DB 721 TGGCGTATCAGATCTGCGAGTCAATAATTTTACATCATATAAATTAATTAATTAATTAAT 780
QY 781 ATCTGAAGCGCTGAAATATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 840
DB 781 ATCTGAAGCGCTGAAATATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 840
QY 841 AAATGAAACCGTGAATGAACGCACTGAAACACAGATATCAAAAGTTTGTGACTTTGAA 900
DB 841 AAATGAAACCGTGAATGAACGCACTGAAACACAGATATCAAAAGTTTGTGACTTTGAA 900
QY 901 GTGCAACGATGATGATGAACATCAGACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 960
DB 901 GTGCAACGATGATGATGAACATCAGACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 960
QY 961 CCGAGGTCATTTTGGCTTTAGGTTGGTCTCAGCGCTTGTGATGTTTGGAGCATAGGTTGCA 1020
DB 961 CCGAGGTCATTTTGGCTTTAGGTTGGTCTCAGCGCTTGTGATGTTTGGAGCATAGGTTGCA 1020

```

```
QY 1021 TTCCTATTGAATATTACCTTGGTTTTCAGAGTCTTTCAGACTCATGATAGTAAAGAGCACC 1080
DB 1021 TTCCTATTGAATATTACCTTGGTTTTCAGAGTCTTTCAGACTCATGATAGTAAAGAGCACC 1080
QY 1081 TGGCAATGATGGAACGAATATTAGGACCCATACCAACACATGATTCAGAAAACCAAGAA 1140
DB 1081 TGGCAATGATGGAACGAATATTAGGACCCATACCAACACATGATTCAGAAAACCAAGAA 1140
QY 1141 AACGCAAGTATTTTCCACATTAACCAAGCTAGATTTGGGATGAAACACATGCTCTGCTGATAG 1200
DB 1141 AACGCAAGTATTTTCCACATTAACCAAGCTAGATTTGGGATGAAACACATGCTCTGCTGATAG 1200
QY 1201 ATGTTAGAGAGCGCTGCAACCGTTGAAGGAATTTATGCTTTGTCATGATGAAGAACATG 1260
DB 1201 ATGTTAGAGAGCGCTGCAACCGTTGAAGGAATTTATGCTTTGTCATGATGAAGAACATG 1260
QY 1261 AGAAACTGTTTGACCTGGTTCGAAGAAATGTTAGAAATATGATCCAACTCAAGAAATTAACCT 1320
DB 1261 AGAAACTGTTTGACCTGGTTCGAAGAAATGTTAGAAATATGATCCAACTCAAGAAATTAACCT 1320
QY 1321 TGGATGAAGCAATGCGACATCCTTTCTTTTGACTTATTAAAAAAGAAATGAAATGGGAATC 1380
DB 1321 TGGATGAAGCAATGCGACATCCTTTCTTTTGACTTATTAAAAAAGAAATGAAATGGGAATC 1380
QY 1381 AGTGTCTTACTATATATCTCTCTAGAGAGATTTACTTTAAGACTGTGTCACTCAACTTAA 1440
DB 1381 AGTGTCTTACTATATATCTCTCTAGAGAGATTTACTTTAAGACTGTGTCACTCAACTTAA 1440
QY 1441 CATTCCTAATATTTTGTAAACATTAATATTTTGTACAGTTAAGTGTAAATATTGTTATG 1500
DB 1441 CATTCCTAATATTTTGTAAACATTAATATTTTGTACAGTTAAGTGTAAATATTGTTATG 1500
QY 1501 TTTTGTATCAATAGCATAAATTAATCTTGTAAAGCAAGTATGCTTTGATAATGCAATGAA 1560
DB 1501 TTTTGTATCAATAGCATAAATTAATCTTGTAAAGCAAGTATGCTTTGATAATGCAATGAA 1560
QY 1561 AAAATTTAAATTTTCTTTTGTAAATTTCAATTTTAAATACCTTTGAAATATCCTT 1620
DB 1561 AAAATTTAAATTTTCTTTTGTAAATTTCAATTTTAAATACCTTTGAAATATCCTT 1620
QY 1621 TGTGTCAGTGTAAATCTGATGATCTTGTCTTTGTGACATGAGGTCACCTCTGAAGT 1680
DB 1621 TGTGTCAGTGTAAATCTGATGATCTTGTCTTTGTGACATGAGGTCACCTCTGAAGT 1680
QY 1681 GATTTTTTTTGGTAAAGGAAATCTTGACTACTTTTATTTCTTAAAGCAATATTTCTTTA 1740
DB 1681 GATTTTTTTTGGTAAAGGAAATCTTGACTACTTTTATTTCTTAAAGCAATATTTCTTTA 1740
QY 1741 TATACTTCAAAATTTAGAACTTAACTTTTAAAGTTTCTTCTGTAAATTTGTTGAACGGGTG 1800
DB 1741 TATACTTCAAAATTTAGAACTTAACTTTTAAAGTTTCTTCTGTAAATTTGTTGAACGGGTG 1800
QY 1801 ATTATTTATTAATCTAGATAGCAGGTACTAGAACCAAACTCAGAAAATGTTTACTGT 1860
DB 1801 ATTATTTATTAATCTAGATAGCAGGTACTAGAACCAAACTCAGAAAATGTTTACTGT 1860
QY 1861 TAGAATCTTATTAATTTTAAAGTGTGTATCTTTTCAATTTGGGTGATGTCAAGGTGATA 1920
DB 1861 TAGAATCTTATTAATTTTAAAGTGTGTATCTTTTCAATTTGGGTGATGTCAAGGTGATA 1920
QY 1921 ACCAGACATTCATGAAAGGCATGAGTTGTCATTTGACAGTTGTTTAAATAAAACC 1980
DB 1921 ACCAGACATTCATGAAAGGCATGAGTTGTCATTTGACAGTTGTTTAAATAAAACC 1980
QY 1981 ACATACACATTTATTTAAGATTTAAATCTTAACCTGGAAGTCAGCTTCGAAAATGACAT 2040
DB 1981 ACATACACATTTATTTAAGATTTAAATCTTAACCTGGAAGTCAGCTTCGAAAATGACAT 2040
QY 2041 TTCCAAGTATGTTGGTGAAGTCAAGATATATAAATAGAAATTTCTGATGAGAGGTTTCAG 2100
DB 2041 TTCCAAGTATGTTGGTGAAGTCAAGATATATAAATAGAAATTTCTGATGAGAGGTTTCAG 2100
QY 2101 TTTTTPAATACCAAGTCTTTAGAGTCTTTAACAATTTGGCCAGCATCTGTTTATCAAAATGACA 2160
```

```
DB 2101 TTTTTPAATACCAAGTCTTTAGAGTCTTTAACAATTTGGCCAGCATCTGTTTATCAAAATGACA 2160
QY 2161 TAAATACGTAACCTATAAGAAATTAAGTTTATTAATTAAGGCAATTTATGCTGTGATTAAT 2220
DB 2161 TAAATACGTAACCTATAAGAAATTAAGTTTATTAATTAAGGCAATTTATGCTGTGATTAAT 2220
QY 2221 TCTTACGGGAGAAAGAGATTTGATTTGGAAGACGATTTGGGAAGAAAGTGTCTGCTGAAAT 2280
DB 2221 TCTTACGGGAGAAAGAGATTTGATTTGGAAGACGATTTGGGAAGAAAGTGTCTGCTGAAAT 2280
QY 2281 TTCCAGAAATTTAATTTGATTTGCTTACATAAATCTTTTTCGACTTCAGAAAAAATAAAAA 2340
DB 2281 TTCCAGAAATTTAATTTGATTTGCTTACATAAATCTTTTTCGACTTCAGAAAAAATAAAAA 2340
QY 2341 AACAAAAAATAAAC 2354
DB 2341 AACAAAAAATAAAC 2354

RESULT 4
US-10-801-671-1
; Sequence 1, Application US/10801671
; Publication No. US20040152123A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000758DIV-III
; CURRENT APPLICATION NUMBER: US/10/801,671
; PRIOR FILING DATE: 2004-03-17
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ. ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-801-671-1

Query Match 100.0%; Score 2354; DB 18; Length 2354;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCAGCTGGGGTTACTTTAAAAAACAATGCTCCATGTCATCCCTCTTGAAGCTTCGCACT 60
DB 1 GCCAGCTGGGGTTACTTTAAAAAACAATGCTCCATGTCATCCCTCTTGAAGCTTCGCACT 60
QY 61 CTGTTGAAGAGGACACTCATCCAGTCAATTTATTTAGAGCAAGGTCCTTTGAATGAGCGAG 120
DB 61 CTGTTGAAGAGGACACTCATCCAGTCAATTTATTTAGAGCAAGGTCCTTTGAATGAGCGAG 120
QY 121 ATTATCGGGACCGGAGATACCTTGACCAATACAGGAATGACTACTGTGAGGATATGTTTC 180
DB 121 ATTATCGGGACCGGAGATACCTTGACCAATACAGGAATGACTACTGTGAGGATATGTTTC 180
QY 181 CTAGACATTTATCAGAGAGACATTTGAAAGCGGGTATCGAATCCACTGCAAGTAAATCTTTCAG 240
DB 181 CTAGACATTTATCAGAGAGACATTTGAAAGCGGGTATCGAATCCACTGCAAGTAAATCTTTCAG 240
QY 241 TCCGACGAGGAGAGCAGTCTTAAAGGAGCGCAATAGACACTGTTCAAGTCAATCACT 300
DB 241 TCCGACGAGGAGAGCAGTCTTAAAGGAGCGCAATAGACACTGTTCAAGTCAATCACT 300
QY 301 CACGTTTGAAGAGGCCACCGAAGGAAAAAGATCCAGGAGTATAGAGGATGATGAGGAGGTC 360
DB 301 CACGTTTGAAGAGGCCACCGAAGGAAAAAGATCCAGGAGTATAGAGGATGATGAGGAGGTC 360
QY 361 ACCTGATCTGTCAAAGTGGAGACGTTTCTTAAGAGCAAGATATGAAATCGTGGACACTTTGG 420
```

Db 361 ACCTGATCTGCAAGTGGAGAGCTTCTAAGAGCAAGATATGAATCGTGGACACTTTGG 420
 Qy 421 GTGAGGAGCCCTTTGGCAAGTTGTAGAGTGCAATGATGATGCGCATGATGCAATGCAATG 480
 Db 421 GTGAGGAGCCCTTTGGCAAGTTGTAGAGTGCAATGATGATGCGCATGATGCAATGCAATG 480
 Qy 481 TAGCAGTGAATATCGTAAAGATGAGCCGTTACCGTGAAGCAGCTGCTTTCAGAAATCC 540
 Db 481 TAGCAGTGAATATCGTAAAGATGAGCCGTTACCGTGAAGCAGCTGCTTTCAGAAATCC 540
 Qy 541 AAGTATTAGACACTTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
 Db 541 AAGTATTAGACACTTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
 Qy 601 TAGAATGCTTTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
 Db 601 TAGAATGCTTTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
 Qy 661 CTTCAGATTTCAATTAAGAAAAACAGCTTTCTGCCATTTCAAAATGACCAATCAGGCAGA 720
 Db 661 CTTCAGATTTCAATTAAGAAAAACAGCTTTCTGCCATTTCAAAATGACCAATCAGGCAGA 720
 Qy 721 TGGCGTATCAGATGCGCAGTCAATTAATTTTTCATCATTAATTAATTAATTAATTAATTA 780
 Db 721 TGGCGTATCAGATGCGCAGTCAATTAATTTTTCATCATTAATTAATTAATTAATTAATTA 780
 Qy 781 ATCTGAAGCCGTAATATTTTGTGTAAGTCTGACTATGATGATGATGATGATGATGATGATG 840
 Db 781 ATCTGAAGCCGTAATATTTTGTGTAAGTCTGACTATGATGATGATGATGATGATGATGATG 840
 Qy 841 AAATGAACGCTGATGAACGACACACTGAAACACACAGATATCAAAAGTTGTTGACCTTGGAA 900
 Db 841 AAATGAACGCTGATGAACGACACACTGAAACACACAGATATCAAAAGTTGTTGACCTTGGAA 900
 Qy 901 GTGCAACGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
 Db 901 GTGCAACGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
 Qy 961 CCGAGGTCATTTGGCTTGAAGTGGTCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTG 1020
 Db 961 CCGAGGTCATTTGGCTTGAAGTGGTCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTG 1020
 Qy 1021 TTCTATTGAATATTTACCTTGGTTTTCAGCTGCTTTCAGCTGCTTTCAGCTGCTTTCAGCT 1080
 Db 1021 TTCTATTGAATATTTACCTTGGTTTTCAGCTGCTTTCAGCTGCTTTCAGCTGCTTTCAGCT 1080
 Qy 1081 TGGCAATGATGGAACGATATTTAGAACCCATPACCAACACATGATGATGATGATGATGATGAT 1140
 Db 1081 TGGCAATGATGGAACGATATTTAGAACCCATPACCAACACATGATGATGATGATGATGATGAT 1140
 Qy 1141 AACGCAAGTATTTTACCATAACAGCTAGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 Db 1141 AACGCAAGTATTTTACCATAACAGCTAGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 Qy 1201 ATGTTAGGAGAGCTGCAACCGTTGAAGGATTTATGCTTTGTCATGATGATGATGATGATGAT 1260
 Db 1201 ATGTTAGGAGAGCTGCAACCGTTGAAGGATTTATGCTTTGTCATGATGATGATGATGATGAT 1260
 Qy 1261 AGAACTGTTTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
 Db 1261 AGAACTGTTTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
 Qy 1321 TGGATGAGCAATTCAGCATCTTTCTTGTGATTTATTAAGAAAGAAATGAAATGGAATC 1380
 Db 1321 TGGATGAGCAATTCAGCATCTTTCTTGTGATTTATTAAGAAAGAAATGAAATGGAATC 1380
 Qy 1381 AGTGGTCTTACTATATCTCTCTAGAGAGATTTACTTAAGACTGCTGCTGCTGCTGCTGCTGCT 1440
 Db 1381 AGTGGTCTTACTATATCTCTCTAGAGAGATTTACTTAAGACTGCTGCTGCTGCTGCTGCTGCT 1440
 Qy 1441 CATCTTAATTTTTTGAACATTAATTTTGTACAGTTAAGTGAATATTTGTATG 1500

Db 1441 CATCTTAATTTTTTGTAAACATTAATTTTGTACAGTTAAGTGAATATTTGTATG 1500
 Qy 1501 TTTTGTATCAATAGCATAAATTAATCTTAAAGCAAGTATGGTCTTGTATGATGATGATGATG 1560
 Db 1501 TTTTGTATCAATAGCATAAATTAATCTTAAAGCAAGTATGGTCTTGTATGATGATGATGATG 1560
 Qy 1561 AAATTAATAATTTAAATTTTCTTTTGAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1620
 Db 1561 AAATTAATAATTTAAATTTTCTTTTGAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1620
 Qy 1621 TGTGTCAGTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680
 Db 1621 TGTGTCAGTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680
 Qy 1681 GATTTTTTTTGAAGTGAAGCAATCTTGAATCTTGAATCTTGAATCTTGAATCTTGAATCTTGA 1740
 Db 1681 GATTTTTTTTGAAGTGAAGCAATCTTGAATCTTGAATCTTGAATCTTGAATCTTGAATCTTGA 1740
 Qy 1741 TATACTTTCAAAATTTAGAACTTAACTTTTAAAGTTTCTTCTGTAATTTGTTGAACGGGTG 1800
 Db 1741 TATACTTTCAAAATTTAGAACTTAACTTTTAAAGTTTCTTCTGTAATTTGTTGAACGGGTG 1800
 Qy 1801 ATTATTAATTAATCTAGATAAGCAGTACTAGAAACCAAACTCAGAAATGTTTACTGT 1860
 Db 1801 ATTATTAATTAATCTAGATAAGCAGTACTAGAAACCAAACTCAGAAATGTTTACTGT 1860
 Qy 1861 TAGAATTTCTAATTAATTTTAAAGTTTGAATCTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1920
 Db 1861 TAGAATTTCTAATTAATTTTAAAGTTTGAATCTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1920
 Qy 1921 ACCAGACATTTCAAGAAAGCAGTCTTGTCTTGTCTGATGATGATGATGATGATGATGATGATG 1980
 Db 1921 ACCAGACATTTCAAGAAAGCAGTCTTGTCTTGTCTGATGATGATGATGATGATGATGATGATG 1980
 Qy 1981 ACATACACATTTTAAAGATTAATTAATCTTAACTGGAAGTCTGAGTTGGAAAAATGGACAT 2040
 Db 1981 ACATACACATTTTAAAGATTAATTAATCTTAACTGGAAGTCTGAGTTGGAAAAATGGACAT 2040
 Qy 2041 TTCCAGATGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2100
 Db 2041 TTCCAGATGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2100
 Qy 2101 TTTTAAATCAACAGTCTTAAAGTCTTAACTTGGCCAGCATCTGTTTATCAAAATGACA 2160
 Db 2101 TTTTAAATCAACAGTCTTAAAGTCTTAACTTGGCCAGCATCTGTTTATCAAAATGACA 2160
 Qy 2161 TAAATAGTGAACCTTAAAGATTAATTAATTTAGGCAATTTATGCTGCTGATGAT 2220
 Db 2161 TAAATAGTGAACCTTAAAGATTAATTAATTTAGGCAATTTATGCTGCTGATGAT 2220
 Qy 2221 TCTTACGGGAGAAAGAGGATTTGATTTGGAAGCAGTTTGGGAAGAAAGTCTGCTGAAAT 2280
 Db 2221 TCTTACGGGAGAAAGAGGATTTGATTTGGAAGCAGTTTGGGAAGAAAGTCTGCTGAAAT 2280
 Qy 2281 TTTCCAGAAATTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
 Db 2281 TTTCCAGAAATTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
 Qy 2341 AACAAAAAATAAAC 2354
 Db 2341 AACAAAAAATAAAC 2354

RESULT 5
 US-10-425-114-26852
 ; Sequence 26852, Application US/10425114
 ; Publication No. US2004003488A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E

APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26852
; LENGTH: 4035
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4654-056-G4_FLI
US-10-425-114-26852

Query Match		81.1%;	Score 1908.4;	DB 17;	Length 4035;
Best Local Similarity		99.9%;	Pred. No. 0;		
Matches 1909;		Conservative	0;	Mismatches	1; Indels 0; Gaps 0;
Qy	398	ATATGAATCGTGGACACTTTGGGTGAAGAGCGCTTTGGCAAAGTTGTAGAGTGCATTGA	457		
Db	2126	AGATGAATCGTGGACACTTTGGGTGAAGAGCGCTTTGGCAAAGTTGTAGAGTGCATTGA	2185		
Qy	458	TCATGGCATGATGGCATGCAATGTAGCAGTGAATAATCGTAAAAATGTAGGCCGTTACCG	517		
Db	2186	TCATGGCATGATGGCATGCAATGTAGCAGTGAATAATCGTAAAAATGTAGGCCGTTACCG	2245		
Qy	518	TGAAGCAGCTCGTTCAGAAATCCCAAGTATTAGACACATTAATAGTACTGATCCCAATAG	577		
Db	2246	TGAAGCAGCTCGTTCAGAAATCCCAAGTATTAGACACATTAATAGTACTGATCCCAATAG	2305		
Qy	578	TGCTCTCCGATGTGTCAGATGCTAGAAATGGTTTGATCATCATGTCATGTTGTGATTGT	637		
Db	2306	TGCTCTCCGATGTGTCAGATGCTAGAAATGGTTTGATCATCATGTCATGTTGTGATTGT	2365		
Qy	638	GTTTGAATCTAGGCACTTAGTACTTACGATTTCAATTAAGAAACAGCTTTCTGCCATT	697		
Db	2366	GTTTGAATCTAGGCACTTAGTACTTACGATTTCAATTAAGAAACAGCTTTCTGCCATT	2425		
Qy	698	TCAATTCACACATCAGCGAGATGGCGTATCAGATCTGCCAGTCAATTAATTTTACATA	757		
Db	2426	TCAATTCACACATCAGCGAGATGGCGTATCAGATCTGCCAGTCAATTAATTTTACATA	2485		
Qy	758	TCATAATAAATTAACCCATACAGATCTGAAGCGCTGAAATATTTTGTGTGAAGTCTGA	817		
Db	2486	TCATAATAAATTAACCCATACAGATCTGAAGCGCTGAAATATTTTGTGTGAAGTCTGA	2545		
Qy	818	CTATGTAGTCAAAATATAATTTCTAAATGAACGTGATGAACGCACACTGAAAAACACAGA	877		
Db	2546	CTATGTAGTCAAAATATAATTTCTAAATGAACGTGATGAACGCACACTGAAAAACACAGA	2605		
Qy	878	TATCAAAAGTTGTGACTTTGGAAGTGCACGTATGATGATGAACATCAGTACTTTGGT	937		
Db	2606	TATCAAAAGTTGTGACTTTGGAAGTGCACGTATGATGATGAACATCAGTACTTTGGT	2665		
Qy	938	GTCTACCCGGCACTACAGAGCTCCGAGGTCAATTTGGCTTTAGGTGGTCTCAGCCCTTG	997		
Db	2666	GTCTACCCGGCACTACAGAGCTCCGAGGTCAATTTGGCTTTAGGTGGTCTCAGCCCTTG	2725		
Qy	998	TGATGTTTGGAGCATAGGTTGCATTTCTTATTGAATATTACCTTGGTTTTCACAGTCTTTCA	1057		
Db	2726	TGATGTTTGGAGCATAGGTTGCATTTCTTATTGAATATTACCTTGGTTTTCACAGTCTTTCA	2785		
Qy	1058	GACTCATGATAGTAAGACACCTGGCAATGATGAACGAAATATTAGACCCATACACACA	1117		
Db	2786	GACTCATGATAGTAAGACACCTGGCAATGATGAACGAAATATTAGACCCATACACACA	2845		
Qy	1118	ACACATGATTACAGAAAACAAAGAAAACGCAAGTATTTTCCACATACACAGCTAGATTGGGA	1177		
Db	2846	ACACATGATTACAGAAAACAAAGAAAACGCAAGTATTTTCCACATACACAGCTAGATTGGGA	2905		
Qy	1178	TGAACACAGTTCTGCTGCTAGATATGTTAGGAGACGCTGCAAAACCGTTGAAGGAATTTAT	1237		

Db	2906	TGAACACAGTTCTGCTGCTAGATATGTTAGGAGACGCTGCAAAACCGTTGAAGGAATTTAT	2965		
Qy	1238	GCCTTTGTCAATGATGAAGAACATAGAAAACCTGTTTGACCTGTTGCGAAGAAATTTAGAATA	1297		
Db	2966	GCCTTTGTCAATGATGAAGAACATAGAAAACCTGTTTGACCTGTTGCGAAGAAATTTAGAATA	3025		
Qy	1298	TGATCCAACTCAAGAAATTTACCTTGGATGAAGCAATTCAGCATCTTTCTTTGACTTATTT	1357		
Db	3026	TGATCCAACTCAAGAAATTTACCTTGGATGAAGCAATTCAGCATCTTTCTTTGACTTATTT	3085		
Qy	1358	AAAAAGAAATGAATCGGAATCAGTGTCTTACTATATATCTCTCTAGAAGAGATTTACT	1417		
Db	3086	AAAAAGAAATGAATCGGAATCAGTGTCTTACTATATATCTCTCTAGAAGAGATTTACT	3145		
Qy	1418	TAAGACTGTGTCACTCAACTAAACATTTCTAATATTTTGTATAAATTTTGTATAAATTTTGTATA	1477		
Db	3146	TAAGACTGTGTCACTCAACTAAACATTTCTAATATTTTGTATAAATTTTGTATAAATTTTGTATA	3205		
Qy	1478	CAGTTAAGTGAATAATTTGTATGTTTGTATCAATAGCATTAATTTAACTTGTTAAGCAAGT	1537		
Db	3206	CAGTTAAGTGAATAATTTGTATGTTTGTATCAATAGCATTAATTTAACTTGTTAAGCAAGT	3265		
Qy	1538	ATGCTCTTGTGAATGCAATTTAGAAAATTTAAAAATTTTCTTTTGTAAAAATTTACCATTT	1597		
Db	3266	ATGCTCTTGTGAATGCAATTTAGAAAATTTAAAAATTTTCTTTTGTAAAAATTTACCATTT	3325		
Qy	1598	TTAAATACCTTTGAAATATCTTTGTGTCAGTGAATAATTTAACTTGTGATTTGCTTTTGT	1657		
Db	3326	TTAAATACCTTTGAAATATCTTTGTGTCAGTGAATAATTTAACTTGTGATTTGCTTTTGT	3385		
Qy	1658	TACATGAGGTGACCTCTGAAGTGAATTTTGTGAGTAAAAAGGAAATCTTGACTACTTTTA	1717		
Db	3386	TACATGAGGTGACCTCTGAAGTGAATTTTGTGAGTAAAAAGGAAATCTTGACTACTTTTA	3445		
Qy	1718	TATTTCTTAAAGGAATATTTCTTTATATCTTCAAAATTTAGAACTTTAACTTTAAAAAGTTT	1777		
Db	3446	TATTTCTTAAAGGAATATTTCTTTATATCTTCAAAATTTAGAACTTTAACTTTAAAAAGTTT	3505		
Qy	1778	CTTCTGTAATTTGTTGAAACGGGTGATTTATTAATTAATCTCTAGATAAGCAGGTACTAGAAC	1837		
Db	3506	CTTCTGTAATTTGTTGAAACGGGTGATTTATTAATTAATCTCTAGATAAGCAGGTACTAGAAC	3565		
Qy	1838	AAAACTCAGAAAATTTGTTTACCTGTAGAAATTTCTAATTTTAAAGTGTGTTCTTTT	1897		
Db	3566	AAAACTCAGAAAATTTGTTTACCTGTAGAAATTTCTAATTTTAAAGTGTGTTCTTTT	3625		
Qy	1898	CATTGGGTGATGTCAGGGTGATAACAGACATTCATGGAAGGCAATGCAAGTTTGTCCATT	1957		
Db	3626	CATTGGGTGATGTCAGGGTGATAACAGACATTCATGGAAGGCAATGCAAGTTTGTCCATT	3685		
Qy	1958	GTGACAGTTTCTTTTAAATAAAACCAACATACACACTTTTAAAGATTTAAATCTAACTGGA	2017		
Db	3686	GTGACAGTTTCTTTTAAATAAAACCAACATACACACTTTTAAAGATTTAAATCTAACTGGA	3745		
Qy	2018	AAGTCAGCTTGGAAAAATTTGGAATTTCCAAATGTTTGGTGGTCAAGATATAAAAAATA	2077		
Db	3746	AAGTCAGCTTGGAAAAATTTGGAATTTCCAAATGTTTGGTGGTCAAGATATAAAAAATA	3805		
Qy	2078	GAATTTCTGATGAGAGTTTTCAGTTTAAATACCAAGTCTTACCAAGTCTTAAACATTCGC	2137		
Db	3806	GAATTTCTGATGAGAGTTTTCAGTTTAAATACCAAGTCTTACCAAGTCTTAAACATTCGC	3865		
Qy	2138	CAGCATCTGTTTATCAAAATGACATAAATTCCTAAACCTTATTAAGAAATTAAGTTTAAAT	2197		
Db	3866	CAGCATCTGTTTATCAAAATGACATAAATTCCTAAACCTTATTAAGAAATTAAGTTTAAAT	3925		
Qy	2198	AGGCAATTTATGCTGTGATTAATTTCTTACGGGAGAAAGAGATTTGATTTGGAAGCAGTT	2257		
Db	3926	AGGCAATTTATGCTGTGATTAATTTCTTACGGGAGAAAGAGATTTGATTTGGAAGCAGTT	3985		
Qy	2258	TGGGAAGAAAGTCTGCTGAAATTTCCAGAAATTTAAATTTGATTTGGTTACAT	2307		
Db	3986	TGGGAAGAAAGTCTGCTGAAATTTCCAGAAATTTAAATTTGATTTGGTTACAT	4035		

```

RESULT 6
US-10-267-502-135
; Sequence 135, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 135
; LENGTH: 1446
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-267-502-135

Query Match      54.6%; Score 1286; DB 17; Length 1446;
Best Local Similarity 100.0%; Pred. No. 1.7e-255;
Matches 1286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85  GTCATTATTAGAAGCAAGGTCCTTGAATGAGCGAGATTATCGGACCGGAGATACGTTG 144
Db 161  GTCATTATTAGAAGCAAGGTCCTTGAATGAGCGAGATTATCGGACCGGAGATACGTTG 220

Qy 145  ACGAATACAGGAATGACTACTGTGAGGATATGTTCTAGACATTATCACAGACATTTG 204
Db 221  ACGAATACAGGAATGACTACTGTGAGGATATGTTCTAGACATTATCACAGACATTTG 280

Qy 205  AAAGCGGTATCGAATCCACATGCGAGTAATCTTCACTCCGACGAGGAGAGAGTCTTA 264
Db 281  AAAGCGGTATCGAATCCACATGCGAGTAATCTTCACTCCGACGAGGAGAGAGTCTTA 340

Qy 265  AAAGGAAGCGCAATAGACACATGTTCAAGTCATCATGTCACGTTGGAAGAGCCACCGAAGGA 324
Db 341  AAAGGAAGCGCAATAGACACATGTTCAAGTCATCATGTCACGTTGGAAGAGCCACCGAAGGA 400

Qy 325  AAAGATCCAGAGATATAGAGATGATGAGGAGGTCACCTGATCTGTCAAAAGTGAGAGC 384
Db 401  AAAGATCCAGAGATATAGAGATGATGAGGAGGTCACCTGATCTGTCAAAAGTGAGAGC 460

Qy 385  TTCTAAGAGCAAGATATGAATCGTGGACACTTTGGGTGAAGGAGCCTTTGGCAAGTTG 444
Db 461  TTCTAAGAGCAAGATATGAATCGTGGACACTTTGGGTGAAGGAGCCTTTGGCAAGTTG 520

Qy 445  TAGAGTGCAATGATCATGGCATGGATGGCATGTCATGTAGCAGTGAAATCGTAAAAAATG 504
Db 521  TAGAGTGCAATGATCATGGCATGGATGGCATGTCATGTAGCAGTGAAATCGTAAAAAATG 580

Qy 505  TAGCCCGTTACCGTGAAGCAGCTGTTCCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTA 564
Db 581  TAGCCCGTTACCGTGAAGCAGCTGTTCCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTA 640

Qy 565  CTGATCCCAATAGTGTCTTCGATGTGTCAGATGCTAGAGTGTGTTGATCATCATGTC 624
Db 641  CTGATCCCAATAGTGTCTTCGATGTGTCAGATGCTAGAGTGTGTTGATCATCATGTC 700

Qy 625  ATGTTTGTATTGTTGTAACACTCTGGGACTTAGTACTTACGATTTCAATTAAGAAAAACA 684
Db 701  ATGTTTGTATTGTTGTAACACTCTGGGACTTAGTACTTACGATTTCAATTAAGAAAAACA 760

Qy 685  GCTTTCTGCCATTTCAATTTACACATCAGGAGATGGGATACAGATCTGCCAGTCAA 744
Db 761  GCTTTCTGCCATTTCAATTTACACATCAGGAGATGGGATACAGATCTGCCAGTCAA 820

Qy 745  TAAATTTTATACATCAATAAATTAACCCATACAGATCTGAAGCCTGAAAAATATTTGT 804
Db 821  TAAATTTTATACATCAATAAATTAACCCATACAGATCTGAAGCCTGAAAAATATTTGT 880

```

```

Qy 805  TTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTAAAAAGAAACGCTGATGAACGACAC 864
Db 881  TTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTAAAAAGAAACGCTGATGAACGACAC 940

Qy 865  TGAATAACACAGATATCAAAAGTTGTTGACCTTTGGAAGTGAACGATGATGATGAATCAATC 924
Db 941  TGAATAACACAGATATCAAAAGTTGTTGACCTTTGGAAGTGAACGATGATGATGAATCAATC 1000

Qy 925  ACAGTACTTTGGTGTCTTACCCGGCAGCTACAGAGTCCCGAGGTCATTTGGCTTTAGGTT 984
Db 1001  ACAGTACTTTGGTGTCTTACCCGGCAGCTACAGAGTCCCGAGGTCATTTGGCTTTAGGTT 1060

Qy 985  GGTCTCAGCCCTTGTGATGTTTGGAGCATAGGTTGCTATTCTTATTGAATATTACCTTGGTT 1044
Db 1061  GGTCTCAGCCCTTGTGATGTTTGGAGCATAGGTTGCTATTCTTATTGAATATTACCTTGGTT 1120

Qy 1045  TCACAGTCTTTCAGACTCATGATAGTAAAGAGACACCTGGCAATGATGAACGAAATATTAG 1104
Db 1121  TCACAGTCTTTCAGACTCATGATAGTAAAGAGACACCTGGCAATGATGAACGAAATATTAG 1180

Qy 1105  GACCCATACCAACACATGATTTTCAGAAACAGAAACGCAAGTATTTTCCACCATACCC 1164
Db 1181  GACCCATACCAACACATGATTTTCAGAAACAGAAACGCAAGTATTTTCCACCATACCC 1240

Qy 1165  AGCTAGATTTGGGATGAACACAGTTTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGT 1224
Db 1241  AGCTAGATTTGGGATGAACACAGTTTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGT 1300

Qy 1225  TGAAGGAATTTATGCTTTGTCATGATGAAGAAACATGAGAAACTGTTTGAACCTGGTTGCAA 1284
Db 1301  TGAAGGAATTTATGCTTTGTCATGATGAAGAAACATGAGAAACTGTTTGAACCTGGTTGCAA 1360

Qy 1285  GAATGTTAGAAATGATGTCCAACTCAAGAAATTTACCTTGGATGAAGCAATTGCAGCATCCTT 1344
Db 1361  GAATGTTAGAAATGATGTCCAACTCAAGAAATTTACCTTGGATGAAGCAATTGCAGCATCCTT 1420

Qy 1345  TCTTTGACTTTATTAATAAGAAATGA 1370
Db 1421  TCTTTGACTTTATTAATAAGAAATGA 1446

RESULT 7
US-10-641-643-699
; Sequence 699, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
;             Susan G. Stuart
;             Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
;                     GENE EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSER: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/641,643
; FILING DATE: 14-Aug-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.

```


Db 419 AGAGAAACGTAATAGACCCCTGTGCAAGTCATCATGTCGATTCGAAGAGCCACCGAAGGA 478
Qy 325 AAAGATCCAGAGATATAGAGATGATGAGGAGGTCCACTGATCTGTCAAAGTGAGACG 384
Db 479 AAAGATCCAGAGATATAGAGATGATGAGGAGGTCCACTGATCTGTCAAAGTGAGACG 538
Qy 385 TTCTAAGAGCAAGATATGAAATCGTGACACTTTGGGTGAAGGAGCCCTTTGGCAAAAGTTG 444
Db 539 TTCTAAGAGCAAGATATGAAATCGTGACACTTTAGGTGAAGGAGCCCTTTGGCAAAAGTTG 598
Qy 445 TAGAGTGCAATGATGATGCGCATGAGTGCATGTAGCATGTAGCATGAAATCTGTAATAATG 504
Db 599 TAGAGTGCAATGATGATGCGCATGAGTGCCTTACATGTAGCATGAAATCTGTAATAATG 658
Qy 505 TAGGCGGTTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCAGCTTTAAATAGTA 564
Db 659 TAGGACGTTACCGGAGGAGCAGCTCGTTCGAAATCCAAAGTATTAGAGCAGCTTTGAACAGCA 718
Qy 565 CTGATCCCAATAGTGTCTTCGAGTGTCCAGATGCTAGAAATGGTTGATCATCATGATGTC 624
Db 719 CTGACCCCAACAGTGTCTTCGAGTGTCCAGATGCTAGAGTGGTTGATCATCATGATGTC 778
Qy 625 ATGTTTGTATGTTGTTGAACTA CTGCGGACTTAGTACTTACATGATGATTTATTAAGAAATA 684
Db 779 ATGTTTGTATGTTGTTGAGCTGTGCGGACTTAGTACTTACATGATGATTTATTAAGAAATA 838
Qy 685 GCTTCTGCCATTTCAAAATTCACACATCAGGAGATGGCGTATCAGATCTGCCAGTCAA 744
Db 839 GTTCTGCCATTTCAAAATTCACATCAGGAGAAATGGCTTATCAGATCTGCCAGTCTA 898
Qy 745 TAAATTTTTTACATCATATAATTAATTAACCCATACAGATCTGAAGCCTGAAAATATTTTGT 804
Db 899 TAAATTTTTTACATCATATAATTAATTAACACACGAGCTTAACCTGAAAATATTTTAT 958
Qy 805 TTGTGAAGTCTGACTATGATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 864
Db 959 TTGTGAAGTCTGACTATGATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1018
Qy 865 TGAATAACACAGATATCAAAAGTGTGTAATTTGGAAGTGCACATATAGACGCAACATC 924
Db 1019 TGAATAACACAGATATCAAAAGTGTGTAATTTGGAAGTGCACATATAGACGCAACATC 1078
Qy 925 ACAGTACTTTGGTGTCTACCGGCACTACAGAGTCCCGAGGTCAATTTTGGCTTTAGGTT 984
Db 1079 ATAGTACTTTGGTGTCTCAAGGCACTACAGGCTCCAGAGTCAATTTTGGCTCTAGTT 1138
Qy 985 GGTCTCAGCCTTTGATGTTTGGAGCATAGTGTGATCTTATTTGAATATTAATTAATTAATTAAT 1044
Db 1139 GGTCTCAGCCTTTGATGTTTGGAGCATAGTGTGATCTTATTTGAATATTAATTAATTAATTAAT 1198
Qy 1045 TCACAGTCTTTCAGACTCATATAGTAAGAGCACTCGCATATGATGGAACGAATATTTAG 1104
Db 1199 TCACAGTCTTTCAGACCCACAGATAGTAAGAGCACTCGCATATGATGGAAGCGGATCTTAG 1258
Qy 1105 GACCCATACCAACACACATGATTCAGAAAAACGAAAGTAATTTTCAACATAACC 1164
Db 1259 GACCCATCCAGCAGATATGATCCAGAGACAAGGAAACGCAAGTATTTTCAACATAACC 1318
Qy 1165 AGCTAGATGGGATGAAACACAGTCTTGTGTTAGATATGTTAGGAGCGCTGCAAAACGCT 1224
Db 1319 AGCTAGATGGGACAGCATAGTTCAAGTGGAGATATGTTAGGAGACGCTGCAAGCGCT 1378
Qy 1225 TGAAGGAATTTATGCTTTGATCATGATGAAGACATGAGAAATCTGTTGACCTGTTGCA 1284
Db 1379 TAAAGGAATTTATGCTGTGATGACGAAAGAGCATGAGAAATCTGTTGACCTGTTGCA 1438
Qy 1285 GAAATGTAAGATATGATCCAACTCAAGAAATTAACCTTGATGAAGCATTTGAGCATCTCT 1344
Db 1439 GAAATGTTGAGATGATGACCCAGGAGAGGATCACCTTGGATGAAGCATTTGAGCATCTCT 1498
Qy 1345 TCTTTGACTTTATTAAGAAAGAAATGAAATGG 1375
|||||

Db 1499 TCTTGACTTATTAAAAAGGAATGAGTGGG 1529
RESULT 9
US-10-825-177-26
; Sequence 26, Application US/10825177
; Publication No. US20040259220A1
; GENERAL INFORMATION:
; APPLICANT: ULLRICH, Axel
; APPLICANT: NAVLER, Oliver
; TITLE OF INVENTION: CLK PROTEIN KINASES AND RELATED PRODUCTS AND METHODS
; FILE REFERENCE: 038602/0431
; CURRENT APPLICATION NUMBER: US/10/825,177
; CURRENT FILING DATE: 2004-04-16
; PRIOR APPLICATION NUMBER: US/09/905,999
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 09/127,248
; PRIOR FILING DATE: 1999-07-31
; PRIOR APPLICATION NUMBER: PCT/IB97/00946
; PRIOR FILING DATE: 1997-06-17
; PRIOR APPLICATION NUMBER: US 08/877,150
; PRIOR FILING DATE: 1997-06-17
; PRIOR APPLICATION NUMBER: US 60/034,286
; PRIOR FILING DATE: 1996-12-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 1549
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-825-177-26
Query Match 46.6%; Score 1097.4; DB 18; Length 1549;
Best Local Similarity 90.6%; Pred. No. 1.7e-216;
Matches 1170; Conservative 0; Mismatches 121; Indels 0; Gaps 0;
Qy 85 GTCATATTTAGAACGAAGTCCCTTGAATGAGCGAGATATTCGGGACCGGAGATACGTTG 144
Db 239 GTCATATTTAGAACGAAGTCCCTTGAATGAGCGAGATATTCGGGACCGGAGATACATG 298
Qy 145 ACGNATACAGGATGACTACTGTGAAGGATATCTTCTAGACATATTCAGAGACATTC 204
Db 299 ATGAATACAGAAATGACTACTGCGAAGGATATGTTTCCAGACATATTCATAGAGACGTT 358
Qy 205 AAAGCGGGTATCGAATCCATCGCAGTAAATCTTTCAGTCCGACGAGGAGAGCAGTCCCTA 264
Db 359 AAAGCCTTACCGGATCCATTCAGTAAATCTTTCAGTCCGAGGAGGAGAGCAGCCCTA 418
Qy 265 AAAGGAGCGCAATAGACACTGTTCAGTCAATCAGTCA CGTTTCGAGAGCCACCGAAGGA 324
Db 419 AGAAGAGCGTAATAGACCCCTGTGCAAGTCAATCAGTCGCAATTCGAGAGCCACCGAAGGA 478
Qy 325 AAAGTCCAGGATATAGAGATGATGAGGAGTCACTGATCTGTCAAAGTGGAGACG 384
Db 479 AAAGATCCAGGATATAGAGGATGATGAGGAGGTCACTGATCTGTCAAAGTGGAGACG 538
Qy 385 TTCTAAGAGCAAGATATGAAATCGTGACACTTTGGGTGAAGGAGCCCTTTGGCAAAAGTTG 444
Db 539 TTCTAAGAGCAAGATATGAAATCGTGACACTTTAGTGAAGGAGCCCTTTGGCAAAAGTTG 598
Qy 445 TAGAGTGCAATGATGATGCGCATGAGTGCATGTAGCATGAAATCTGTAATAATG 504
Db 599 TAGAGTGCAATGATGATGCGCATGAGTGCCTTACATGTAGCATGAAATCTGTAATAATG 658
Qy 505 TAGGCGGTTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCAGCTTTAAATAGTA 564
Db 659 TAGGACGTTACCGGAGGAGCAGCTCGTTCGAAATCCAAAGTATTAGAGCAGCTTTGAACAGCA 718
Qy 565 CTGATCCCAATAGTGTCTTCGAGTGTCCAGATGCTAGAAATGGTTGATCATCATGATGTC 624
Db 719 CTGACCCCAACAGTGTCTTCGAGTGTCCAGATGCTAGAGTGGTTGATCATCATGATGTC 778
Qy 625 ATGTTTGTATGTTGTTGAACTA CTGCGGACTTAGTACTTACATGATGATTTATTAAGAAATA 684
|||||

QY 1165 AGCTAGATTGGATGAACACAGTTCTGCTGGTAGATATGTTAGGAGACGCTGCAACCGT 1224
 Db 1241 AGCTAGATTGGAGAGAGATAGTTAGCTGGGAGATATGTTAGGAGACGCTGCAACCGT 1300
 QY 1225 TGAAGGAATTTATGCTTTGTCATGATGAAGAACATGAGAACTGTTTGAACCTGGTTGCAA 1284
 Db 1301 TAAAGGAATTTATGCTGGTGTATGACGAAGAGCATGAGAACGCTGTTTGACCTGGTTGCAA 1360
 QY 1285 GAATGTTAGATATGATCCAACTCAAGAAATACCTTGGATGAAGCAATGCGAGCATCCTT 1344
 Db 1361 GAATGTTGGAGTATGACCCAGGAGAGGATCACCTTGGATGAAGCAATGCGAGCACCTT 1420
 QY 1345 TCTTTGACTTATTAAGAAAGAAATGA 1370
 Db 1421 TCTTTGACTTATTAAGAAAGAAATGA 1446

RESULT 11
 US-09-810-671-3
 ; Sequence 3, Application US/09810671
 ; Publication No. US20020076783A1
 ; GENERAL INFORMATION:
 ; APPLICANT: YAN, Chunhua et al
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; FILE REFERENCE: CL000758
 ; CURRENT APPLICATION NUMBER: US/09/810,671
 ; CURRENT FILING DATE: 2001-06-08
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 21234
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-810-671-3

Query Match 46.6%; Score 1097; DB 9; Length 21234;
 Best Local Similarity 100.0%; Pred. No. 6.7e-216; Indels 0; Gaps 0;
 Matches 1097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1228 AGGAATTTATGCTTTGTCATGATGAAGAACATGAGAACTGTTTGACCTGGTTGCGAAGAA 1287
 Db 18138 AGGAATTTATGCTTTGTCATGATGAAGAACATGAGAACTGTTTGACCTGGTTGCGAAGAA 18197
 QY 1288 TGTTAGATATGATCCAACTCAAGAAATACCTTGGATGAAGCAATGCGAGCATCCTTTCT 1347
 Db 18198 TGTTAGATATGATCCAACTCAAGAAATACCTTGGATGAAGCAATGCGAGCATCCTTTCT 18257
 QY 1348 TTGACTTATTAAGAAAGAAATGAATGGGAATCAGTGCTTACTATATCTTCTAGTA 1407
 Db 18258 TTGACTTATTAAGAAAGAAATGAATGGGAATCAGTGCTTACTATATCTTCTAGTA 18317
 QY 1408 AGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTAAA 1467
 Db 18318 AGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTAAA 18377
 QY 1468 TTATTTGTACAGTTAAGTGAATATTTGATGTTTGTATCAATAGCATTAATTAACCTG 1527
 Db 18378 TTATTTGTACAGTTAAGTGAATATTTGATGTTTGTATCAATAGCATTAATTAACCTG 18437
 QY 1528 TTAAGCAAGTATGGTCTTGTAAATGCAATTAAGAAATTAATTAATTTTCTTTTGTAA 1587
 Db 18438 TTAAGCAAGTATGGTCTTGTAAATGCAATTAAGAAATTAATTAATTTTCTTTTGTAA 18497
 QY 1588 ATTACCAATTTTAAATACCTTTGAAATATCCTTTGTGCCAGTGATAAATGATGATC 1647
 Db 18498 ATTACCAATTTTAAATACCTTTGAAATATCCTTTGTGCCAGTGATAAATGATGATC 18557
 QY 1648 TTGCCTTTTGTACATGGAGGTCACCTCTGAAGTGAATTTTTTTTGTAGTAAAGGAATCTT 1707
 Db 18558 TTGCCTTTTGTACATGGAGGTCACCTCTGAAGTGAATTTTTTTTGTAGTAAAGGAATCTT 18617

QY 1708 GACTACTTTATATTTCTTAAGGAATATTTCTTTATATATCTTCAAAATTTAGAACTTAACCTT 1767
 Db 18618 GACTACTTTATATTTCTTAAGGAATATTTCTTTATATATCTTCAAAATTTAGAACTTAACCTT 18677
 QY 1768 AAAAGTTTTTCTTCTGTAATTTGTAACCGGTTGATTTATTAATTAACCTAGATTAAGCAGGT 1827
 Db 18678 AAAAGTTTTTCTTCTGTAATTTGTAACCGGTTGATTTATTAATTAACCTAGATTAAGCAGGT 18737
 QY 1828 ACTAGAAACCAAACTCAGAAATGTTTACTGTTAGAAATTTCTATTAATTTTAAAGTGTG 1887
 Db 18738 ACTAGAAACCAAACTCAGAAATGTTTACTGTTAGAAATTTCTATTAATTTTAAAGTGTG 18797
 QY 1888 TATTTCTTTTCAATTTGGTGTATGTCAGGTTGATTAACAGACATTCATCGAAAGCATGTCAG 1947
 Db 18798 TATTTCTTTTCAATTTGGTGTATGTCAGGTTGATTAACAGACATTCATCGAAAGCATGTCAG 18857
 QY 1948 TTTGTCCATTTGTCAGAGTTTGTATTAATAAACACATACACATTTATTTAAAGATTAAAA 2007
 Db 18858 TTTGTCCATTTGTCAGAGTTTGTATTAATAAACACATACACATTTATTTAAAGATTAAAA 18917
 QY 2008 TCTAACTGGAAAGTCAGCTTGGAAATGGACATTTCCAACTATGTTTGGTGAAGTACAGAA 2067
 Db 18918 TCTAACTGGAAAGTCAGCTTGGAAATGGACATTTCCAACTATGTTTGGTGAAGTACAGAA 18977
 QY 2068 TATAAAATAGAAATCTGATGAGAGGTTTCAGTTTTTAAATACCAAGTCTCTTAGGAGTCT 2127
 Db 18978 TATAAAATAGAAATCTGATGAGAGGTTTCAGTTTTTAAATACCAAGTCTCTTAGGAGTCT 19037
 QY 2128 TAACATTTGGCCAGCATCTGTTATCAAAATGACATTAATAGCTAACTAAGAAATTAAG 2187
 Db 19038 TAACATTTGGCCAGCATCTGTTATCAAAATGACATTAATAGCTAACTAAGAAATTAAG 19097
 QY 2188 TTTTATTAATTTAGGCAATTTATGTCGTGATAATTTCTTACGGGAGAAAGAGATTTGATTG 2247
 Db 19098 TTTTATTAATTTAGGCAATTTATGTCGTGATAATTTCTTACGGGAGAAAGAGATTTGATTG 19157
 QY 2248 GAAAGCAGTTTGGGAGAAAGTCTGCTGAAATTTCCAGAAATTTAAATTTGATTGTTACAT 2307
 Db 19158 GAAAGCAGTTTGGGAGAAAGTCTGCTGAAATTTCCAGAAATTTAAATTTGATTGTTACAT 19217
 QY 2308 AAACCTTTTTCAGTTTCAG 2324
 Db 19218 AAACCTTTTTCAGTTTCAG 19234

RESULT 12
 US-10-109-854-3
 ; Sequence 3, Application US/10109854
 ; Publication No. US20020119548A1
 ; GENERAL INFORMATION:
 ; APPLICANT: YAN, Chunhua et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; FILE REFERENCE: CL000758DIV
 ; CURRENT APPLICATION NUMBER: US/10/109,854
 ; CURRENT FILING DATE: 2002-04-01
 ; PRIOR APPLICATION NUMBER: 60/227,470
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: 09/810,671
 ; PRIOR FILING DATE: 2001-03-19
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 21234
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-10-109-854-3

Query Match 46.6%; Score 1097; DB 13; Length 21234;
 Best Local Similarity 100.0%; Pred. No. 6.7e-216; Indels 0; Gaps 0;
 Matches 1097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1228 AGGAATTTATGCTTTGTCAATGATGAAGAACATGAGAAACTGTTTGCACCTGGTTCGAAGAA 1287
Db 18138 AGGAATTTATGCTTTGTCAATGATGAAGAACATGAGAAACTGTTTGCACCTGGTTCGAAGAA 18197
QY 1288 TGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCAATGAGCAATCTTTCT 1347
Db 18198 TGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCAATGAGCAATCTTTCT 18257
QY 1348 TTGACTTTATTAAGAAAGAAATGGAATGGAATCAGTGGTCTTACTATATCTTCTCTAGA 1407
Db 18258 TTGACTTTATTAAGAAAGAAATGGAATGGAATCAGTGGTCTTACTATATCTTCTCTAGA 18317
QY 1408 AGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTAAA 1467
Db 18318 AGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTAAA 18377
QY 1468 TTATTTTGTACAGTTAAGTGTAAATATTTGTATGTTTGTATCAATAGCAATATTAACCTG 1527
Db 18378 TTATTTTGTACAGTTAAGTGTAAATATTTGTATGTTTGTATCAATAGCAATATTAACCTG 18437
QY 1528 TTAACCAAGTATGCTTTGATAATGCATTTAGAAAAATTAATAATTTTCTTTTTCGAA 1587
Db 18438 TTAACCAAGTATGCTTTGATAATGCATTTAGAAAAATTAATAATTTTCTTTTTCGAA 18497
QY 1588 ATTACCATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATAAATGTGATTCATC 1647
Db 18498 ATTACCATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATAAATGTGATTCATC 18557
QY 1648 TTGCTTTTGTACATGGAGGTCACCTCTGAAGTGAATTTTGTAGTAAAGGAATCTTT 1707
Db 18558 TTGCTTTTGTACATGGAGGTCACCTCTGAAGTGAATTTTGTAGTAAAGGAATCTTT 18617
QY 1708 GACTACTTTATATCTTAAAGGAATTTCTTTATATCTTCAAAATTTAGAACTTAACCTTT 1767
Db 18618 GACTACTTTATATCTTAAAGGAATTTCTTTATATCTTCAAAATTTAGAACTTAACCTTT 18677
QY 1768 AAAAGTTTTTCTCTGTAAATTTGTAACGGGTGATTTATTTAACTCTAGATAAGCAGGT 1827
Db 18678 AAAAGTTTTTCTCTGTAAATTTGTAACGGGTGATTTATTTAACTCTAGATAAGCAGGT 18737
QY 1828 ACTAGAAACCAAACTCAGAAAAATGTTTACTGTGTAGAAATCTATTTAAATTTTAACTGTG 1887
Db 18738 ACTAGAAACCAAACTCAGAAAAATGTTTACTGTGTAGAAATCTATTTAAATTTTAACTGTG 18797
QY 1888 TATTTCTTTTCAATCGGGTGATGTCAGGGTGATAACAGACATTTCAATGAAAGGCATGCAG 1947
Db 18798 TATTTCTTTTCAATCGGGTGATGTCAGGGTGATAACAGACATTTCAATGAAAGGCATGCAG 18857
QY 1948 TTTGTCCATTTGTGACAGTTTGTTTAATAAACCACATACACACTTTTATTTAAGATTAAAA 2007
Db 18858 TTTGTCCATTTGTGACAGTTTGTTTAATAAACCACATACACACTTTTATTTAAGATTAAAA 18917
QY 2008 TCTAACTGGAAGTCACTGTGGAATGGAATGGAATTTCCAGATGATGTTGGTGAGTCACAGA 2067
Db 18918 TCTAACTGGAAGTCACTGTGGAATGGAATTTCCAGATGATGTTGGTGAGTCACAGA 18977
QY 2068 TATAAATATGAAATCTGTGATGAGAGGTTTCAGTTTAAATACCAAGTCCCTTAGAGTCT 2127
Db 18978 TATAAATATGAAATCTGTGATGAGAGGTTTCAGTTTAAATACCAAGTCCCTTAGAGTCT 19037
QY 2128 TAAACATTTGCCAGCATCTGTTTATCAAAATGACATAAATACGTAAACCTATTAAGAAATTAAG 2187
Db 19038 TAAACATTTGCCAGCATCTGTTTATCAAAATGACATAAATACGTAAACCTATTAAGAAATTAAG 19097
QY 2188 TTTATTAATAGGCAATTTATGTCGTGTGATATTTCTTACGGGAGAAAGAGGATTTGATG 2247
Db 19098 TTTATTAATAGGCAATTTATGTCGTGTGATATTTCTTACGGGAGAAAGAGGATTTGATG 19157
QY 2248 GAAACGATTTGGGAGAAAGTCTGCTGAAATTTCCAGAAATTTAAATGATTTGGTTACAT 2307
Db 19158 GAAACGATTTGGGAGAAAGTCTGCTGAAATTTCCAGAAATTTAAATGATTTGGTTACAT 19217
```

```
QY 2308 AAACCTTTTGTACTTCAG 2324
Db 19218 AAACCTTTTGTACTTCAG 19234

RESULT 13
US-10-339-656-3
; Sequence 3, Application US/10339656
; Publication No. US20030134319A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; FILE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000758DIV2
; CURRENT APPLICATION NUMBER: US/10/339,656
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/109,854
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3.
; LENGTH: 21234
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-339-656-3
```

```
Query Match 46.6%; Score 1097; DB 15; Length 21234;
Best Local Similarity 100.0%; Pred. No. 6.7e-216;
Matches 1097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1228 AGGAATTTATGCTTTGTCAATGATGAAGAACATGAGAAACTGTTTGCACCTGGTTCGAAGAA 1287
Db 18138 AGGAATTTATGCTTTGTCAATGATGAAGAACATGAGAAACTGTTTGCACCTGGTTCGAAGAA 18197
QY 1288 TGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCAATGAGCAATCTTTCT 1347
Db 18198 TGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCAATGAGCAATCTTTCT 18257
QY 1348 TTGACTTTATTAAGAAAGAAATGGAATGGAATCAGTGGTCTTACTATATCTTCTCTAGA 1407
Db 18258 TTGACTTTATTAAGAAAGAAATGGAATGGAATCAGTGGTCTTACTATATCTTCTCTAGA 18317
QY 1408 AGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTAAA 1467
Db 18318 AGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTAAA 18377
QY 1468 TTATTTTGTACAGTTAAGTGTAAATATTTGTATGTTTGTATCAATAGCAATATTAACCTG 1527
Db 18378 TTATTTTGTACAGTTAAGTGTAAATATTTGTATGTTTGTATCAATAGCAATATTAACCTG 18437
QY 1528 TTAACCAAGTATGCTTTGATAATGCATTTAGAAAAATTAATAATTTTCTTTTTCGAA 1587
Db 18438 TTAACCAAGTATGCTTTGATAATGCATTTAGAAAAATTAATAATTTTCTTTTTCGAA 18497
QY 1588 ATTACCATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATAAATGTGATTCATC 1647
Db 18498 ATTACCATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATAAATGTGATTCATC 18557
QY 1648 TTGCTTTTGTACATGGAGGTCACCTCTGAAGTGAATTTTGTAGTAAAGGAATCTTT 1707
Db 18558 TTGCTTTTGTACATGGAGGTCACCTCTGAAGTGAATTTTGTAGTAAAGGAATCTTT 18617
QY 1708 GACTACTTTATATCTTAAAGGAATTTCTTTATATCTTCAAAATTTAGAACTTAACCTTT 1767
Db 18618 GACTACTTTATATCTTAAAGGAATTTCTTTATATCTTCAAAATTTAGAACTTAACCTTT 18677
QY 1768 AAAAGTTTTTCTCTGTAAATTTGTAACGGGTGATTTATTTAACTCTAGATAAGCAGGT 1827
```

Db 18678 AAAAGTTTTTCTGTGTAATGTTGGAACGGGTGATTAATTAATTAACCTAGATAAGCAGGT 18737
QY 1828 ACTAGAACCAAACTCAGAAAATGTTTACCTGTAGAAATCTTATTAATTTTAAAGTGTG 1887
Db 18738 ACTAGAACCAAACTCAGAAAATGTTTACCTGTAGAAATCTTATTAATTTTAAAGTGTG 18797
QY 1888 TATCTCTTTTTCANTGGGTGATGTCAGGGTGATAACCCAGACATTCATGGAAAGGCATGCAG 1947
Db 18798 TATCTCTTTTTCANTGGGTGATGTCAGGGTGATAACCCAGACATTCATGGAAAGGCATGCAG 18957
QY 1948 TTTGTCCATTTGTGACAGTTTGTGTAATAAAACCACATACACACTTTATTTAAGATTAAAA 2007
Db 18858 TTTGTCCATTTGTGACAGTTTGTGTAATAAAACCACATACACACTTTATTTAAGATTAAAA 18917
QY 2008 TCTAACTGGAAGTCAGCTTCGAAAATGACATTTCCAAAGTATGTTTGGTGAGTCACAGA 2067
Db 18918 TCTAACTGGAAGTCAGCTTCGAAAATGACATTTCCAAAGTATGTTTGGTGAGTCACAGA 18977
QY 2068 TATAAAATAGAAATTCGTATGAGAGGTTTCAGTTTTTAAATACCAAGTCCTTAGGAGTCT 2127
Db 18978 TATAAAATAGAAATTCGTATGAGAGGTTTCAGTTTTTAAATACCAAGTCCTTAGGAGTCT 19037
QY 2128 TAAATTTGCCAGCATCTGTTTATCAAAATGACATAAATACGTAACCTATAGAAATTAAG 2187
Db 19038 TAAATTTGCCAGCATCTGTTTATCAAAATGACATAAATACGTAACCTATAGAAATTAAG 19097
QY 2188 TTTTATTAAATAGGCAATTTATGTCGTGATATTTCTACGGGAGAAAGAGGATTTGATTG 2247
Db 19098 TTTTATTAAATAGGCAATTTATGTCGTGATATTTCTACGGGAGAAAGAGGATTTGATTG 19157
QY 2248 GAAAGCAGTTTGGGAGAAAGTGTCTGCTGAAATTTCCAGAAATTTAAATTTGATTGTTACAT 2307
Db 19158 GAAAGCAGTTTGGGAGAAAGTGTCTGCTGAAATTTCCAGAAATTTAAATTTGATTGTTACAT 19217
QY 2308 AAACCTTTTGTACTTCAG 2324
Db 19218 AAACCTTTTGTACTTCAG 19234

RESULT 14

US-10-801-671-3

; Sequence 3, Application US/10801671

; Publication No. US20040152123A1

; GENERAL INFORMATION:

; APPLICANT: YAN, Chunhua et al.

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: CLO00758DIV-III

; CURRENT APPLICATION NUMBER: US/10/801,671

; PRIOR FILING DATE: 2004-03-17

; PRIOR APPLICATION NUMBER: 60/227,470

; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: 09/810,671

; PRIOR FILING DATE: 2001-03-19

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 21234

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-801-671-3

Query Match 46.6%; Score 1097; DB 18; Length 21234;

Best Local Similarity 100.0%; Pred. No. 6.7e-216;

Matches 1097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1228 AGGAATTTATGCTTTGTCATGATGAAGAACATGAAACTGTTTGACCTGGTTCGAAGAA 1287

Db 18138 AGGAATTTATGCTTTGTCATGATGAAGAACATGAAACTGTTTGACCTGGTTCGAAGAA 18197

QY 1288 TGTTAGAAATGATCCAACTCAAGAAATTAACCTTGATGAAGCATTCGAGCATCTTTCT 1347

|||||

Db 18198 TGTTAGAAATGATCCAACTCAAGAAATTAACCTTGGAATGAAGCATTCGAGCATCTTTCT 18257
QY 1348 TTGACTTTATTAATAAGAAATGAAATGGGAATCAGTGGTCTTACTATATACTTCTCTAGA 1407
Db 18258 TTGACTTTATTAATAAGAAATGAAATGGGAATCAGTGGTCTTACTATATACTTCTCTAGA 18317
QY 1408 AGAGATTACTTTAAGACTGTGTGTCAGTCAACTAAACATTTCTAATA TTTTGTGTAACATTTAA 1467
Db 18318 AGAGATTACTTTAAGACTGTGTGTCAGTCAACTAAACATTTCTAATA TTTTGTGTAACATTTAA 18377
QY 1468 TTAATTTGTACAGTTAAGTAAATATTTGATGTTTGTATCAATAGCATTAATTAACCTTG 1527
Db 18378 TTAATTTGTACAGTTAAGTAAATATTTGATGTTTGTATCAATAGCATTAATTAACCTTG 18437
QY 1528 TTAAGCAAGTATGTCCTGTGATAATGCAATAGAAAAATTAATAATTTTCTTTTGTGAA 1587
Db 18438 TTAAGCAAGTATGTCCTGTGATAATGCAATAGAAAAATTAATAATTTTCTTTTGTGAA 18497
QY 1588 ATTAACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATTAATGATGATGATC 1647
Db 18498 ATTAACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATTAATGATGATGATC 18557
QY 1648 TTGCTTTTGTACATGAGGTCACCTCTGAAGTCATTTTTCAGTAAAGGAAATCTT 1707
Db 18558 TTGCTTTTGTACATGAGGTCACCTCTGAAGTCATTTTTCAGTAAAGGAAATCTT 18617
QY 1708 GACTACTTTTATATCTTAAAGGAATATCTTTATATATCTTCAAAATTTAGAACTTTAACTTT 1767
Db 18618 GACTACTTTTATATCTTAAAGGAATATCTTTATATATCTTCAAAATTTAGAACTTTAACTTT 18677
QY 1768 AAAAGTTTTTCTCTGTAATTTGTTGAAAGGGTGATTAATTAATCTCTAGATAAGCAGGT 1827
Db 18678 AAAAGTTTTTCTCTGTAATTTGTTGAAAGGGTGATTAATTAATCTCTAGATAAGCAGGT 18737
QY 1828 ACTAGAAACCAAACTCAGAAAATGTTTACTGTTAGAAATCTTATTAATTTTAAAGTGTG 1887
Db 18738 ACTAGAAACCAAACTCAGAAAATGTTTACTGTTAGAAATCTTATTAATTTTAAAGTGTG 18797
QY 1888 TATTCTTTTTCATTTGGGTGATGTCAGGGTGATAACCCAGACATTCATGGAAAGGCATGCAG 1947
Db 18798 TATTCTTTTTCATTTGGGTGATGTCAGGGTGATAACCCAGACATTCATGGAAAGGCATGCAG 18857
QY 1948 TTTGTCCATTTGTGACAGTTTCTTTAATAAAACCCACATACACACTTTATTTAAGATTAAAA 2007
Db 18858 TTTGTCCATTTGTGACAGTTTCTTTAATAAAACCCACATACACACTTTATTTAAGATTAAAA 18917
QY 2008 TCTAACTGGAAGTCAGCTTCGAAAATGAGACATTTCCAAAGTATGTTTGGTGAGTCACAGA 2067
Db 18918 TCTAACTGGAAGTCAGCTTCGAAAATGAGACATTTCCAAAGTATGTTTGGTGAGTCACAGA 18977
QY 2068 TATAAAATAGAAATTCGTATGAGAGGTTTCAGTTTTTAAATACCAAGTCCTTAGGAGTCT 2127
Db 18978 TATAAAATAGAAATTCGTATGAGAGGTTTCAGTTTTTAAATACCAAGTCCTTAGGAGTCT 19037
QY 2128 TAAATTTGCCAGCATCTGTTTATCAAAATGACATAAATACGTAACCTATAGAAATTAAG 2187
Db 19038 TAAATTTGCCAGCATCTGTTTATCAAAATGACATAAATACGTAACCTATAGAAATTAAG 19097
QY 2188 TTTTATTAAATAGGCAATTTATGTCGTGATATTTCTACGGGAGAAAGAGGATTTGATTG 2247
Db 19098 TTTTATTAAATAGGCAATTTATGTCGTGATATTTCTACGGGAGAAAGAGGATTTGATTG 19157
QY 2248 GAAAGCAGTTTGGGAGAAAGTGTCTGCTGAAATTTCCAGAAATTTAAATTTGATTGTTACAT 2307
Db 19158 GAAAGCAGTTTGGGAGAAAGTGTCTGCTGAAATTTCCAGAAATTTAAATTTGATTGTTACAT 19217
QY 2308 AAACCTTTTGTACTTCAG 2324
Db 19218 AAACCTTTTGTACTTCAG 19234

RESULT 15

US-10-425-114-26212

Query Match	41.5%	Score	977.4	DB	17	Length	3040
Best Local Similarity	99.9%	Pred. No.	1.4e-191				
Matches	978	Conservative	0	Mismatches	1	Indels	0
Qy	398	ATATGAATCGTGGACAC	TTTGGGTGAAGAGCGCTTTGGCAAAAGTTGTAGAGTGCATTGA	457			
Db	2062	AGATGAATCGTGGACAC	TTTGGGTGAAGAGCGCTTTGGCAAAAGTTGTAGAGTGCATTGA	2121			
Qy	458	TCATGGCATCGATGGCGATG	CATGTAGCAGTGAATAATCGTAAAAAATCTAGCGCGTTACCG	517			
Db	2122	TCATGGCATCGATGGCGATG	CATGTAGCAGTGAATAATCGTAAAAAATCTAGCGCGTTACCG	2181			
Qy	518	TGAAGCAGCTCGTTTCAGAAA	TCCAAATATAGAGCACTTAAATAGTACTCATCCCAATAG	577			
Db	2182	TGAAGCAGCTCGTTTCAGAAA	TCCAAATATAGAGCACTTAAATAGTACTCATCCCAATAG	2241			
Qy	578	TGTCTTCGATGTCCAGATGCT	AGAAATGGTTTGATCATCATGGTCAATGTTTGTATGT	637			
Db	2242	TGTCTTCGATGTCCAGATGCT	AGAAATGGTTTGATCATCATGGTCAATGTTTGTATGT	2301			
Qy	638	GTTTGAAC	TACTGGGACTTAGTACTACGATTTCAATAAGAAACAGCTTCTCGCAAT	697			
Db	2302	GTTTGAAC	TACTGGGACTTAGTACTACGATTTCAATAAGAAACAGCTTCTCGCAAT	2361			
Qy	698	TCAAATTGAC	CACATCAGGCAGATGGCGTATCAGATCTGCCAGTCAATAAATTTTTTACA	757			
Db	2362	TCAAATTGAC	CACATCAGGCAGATGGCGTATCAGATCTGCCAGTCAATAAATTTTTTACA	2421			
Qy	758	TCATAATAA	TTAAACCATACAGATCTGAAAGCCTGAAAAATATTTGTTGTGAAGTCTGA	817			
Db	2422	TCATAATAA	TTAAACCATACAGATCTGAAAGCCTGAAAAATATTTGTTGTGAAGTCTGA	2481			
Qy	818	CTATGTAGTCAA	ATATAATTTCTAAATGAAACGATGAACGCACACTGAAAAACACAGA	877			
Db	2482	CTATGTAGTCAA	ATATAATTTCTAAATGAAACGATGAACGCACACTGAAAAACACAGA	2541			
Qy	878	TATCAAAGTTGTTGAC	TTTGGAGTGAACGATATGATGAACATCACAAGTACTTTGGT	937			
Db	2542	TATCAAAGTTGTTGAC	TTTGGAGTGAACGATATGATGAACATCACAAGTACTTTGGT	2601			
Qy	938	GTCTACCCGC	CACTACAGAGCTCCGAGGTCAATTTTGGCTTTAGTTGGTCTCAGCCCTG	997			
Db	2602	GTCTACCCGC	CACTACAGAGCTCCGAGGTCAATTTTGGCTTTAGTTGGTCTCAGCCCTG	2661			
Qy	998	TGATGTTTGGAG	CATAGGTTGCAATTCATATGAAATATTAACCTTGGTTTCACAGTCTTTCA	1057			
Db	2662	TGATGTTTGGAG	CATAGGTTGCAATTCATATGAAATATTAACCTTGGTTTCACAGTCTTTCA	2721			
Qy	1058	GACTCATGATAG	TAAAGACACCTGGCAATGATGGAAACGAATATTAGGACCCATACACA	1117			

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2005, 23:50:16 ; Search time 166 Seconds
(without alignments)
1036.797 Million cell updates/sec

Title: US-10-801-671-2

Perfect score: 2410
Sequence: 1 MCIPLEASHSVEDTHPSHY.....QRITLDEALQHPPDLLKKK 445

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2410	100.0	445	5	Aae20170 Human kin
2	2315.5	96.1	481	8	ADI26143 Human pro
3	2315.5	96.1	481	8	ADO08026 Human pol
4	2315.5	96.1	481	8	ADP24394 PRO polyp
5	2294.5	95.2	481	4	Aab69705 Human cel
6	2261.5	93.8	481	8	ADI26141 Human pro
7	2252.5	93.5	481	8	ADO08029 Mouse pol
8	2177.5	90.4	479	2	AAW49914 Mouse CLK
9	2127.5	88.3	499	4	Aab65648 Novel pro
10	2127.5	88.3	499	8	ADI29255 Human MAR
11	2117.5	87.9	451	2	AAY27054 Human pro
12	1893	78.5	484	8	ADO08027 Human pol
13	1887	78.3	484	5	AAU75108 Serine an
14	1887	78.3	484	8	ADH58849 Human CDC
15	1887	78.3	484	8	ADR14001 Human NF-
16	1887	78.3	484	8	ADP24719 PRO polyp
17	1823	75.6	483	8	ADO08030 Mouse pol
18	1808.5	75.0	482	2	AAW49911 Mouse CLK
19	1803	74.8	453	8	ADI26145 Human pro
20	1607	66.7	301	8	ADO08031 Mouse pol
21	1549	64.3	441	7	ADG74657 Human kin
22	1535	63.7	283	3	AAB33778 Human sec
23	1526	63.3	350	3	AAB58831 Breast an
24	1517	62.9	283	3	AAB33777 Human sec
25	1422.5	59.0	498	6	Aae37967 Human kin

ALIGNMENTS

RESULT 1

ID	AAE20170	standard; protein; 445 AA.
XX	AAE20170;	
AC	AAE20170;	
XX	18-JUN-2002 (first entry)	
DT	Human kinase protein.	
DE	Human; kinase protein; enzyme; cytostatic; osteopathic; gene expression;	
XX	Human; kinase protein; enzyme; cytostatic; osteopathic; gene expression;	
KW	colon-moderately differentiated adenocarcinoma; chromosome mapping; gene;	
KW	drug screening; therapeutic; gene therapy; tissue typing; chromosome 5;	
KW	bone osteosarcoma.	
XX	Homo sapiens.	
OS	Homo sapiens.	
XX	Key	Location/Qualifiers
FH	Modified-site	10. .13
FT	Modified-site	/note= "Casein kinase II phosphorylation site"
FT	Modified-site	24. .31
FT	Modified-site	/note= "Tyrosine kinase phosphorylation site"
FT	Modified-site	25. .28
FT	Modified-site	/note= "Casein kinase II phosphorylation site"
FT	Modified-site	29. .36
FT	Modified-site	/note= "Tyrosine kinase phosphorylation site"
FT	Modified-site	55. .61
FT	Modified-site	/note= "Tyrosine kinase phosphorylation site"
FT	Modified-site	69. .71
FT	Modified-site	/note= "Protein kinase C phosphorylation site"
FT	Modified-site	72. .74
FT	Modified-site	/note= "Protein kinase C phosphorylation site"
FT	Modified-site	73. .76
FT	Modified-site	/note= "cAMP and cGMP-dependent protein kinase phosphorylation site"
FT	Modified-site	76. .78
FT	Modified-site	/note= "Protein kinase C phosphorylation site"
FT	Modified-site	94. .96
FT	Modified-site	/note= "Protein kinase C phosphorylation site"
FT	Modified-site	97. .100
FT	Modified-site	/note= "cAMP and cGMP-dependent protein kinase phosphorylation site"
FT	Modified-site	102. .105
FT	Modified-site	/note= "Casein kinase II phosphorylation site"
FT	Modified-site	128. .131
FT	Modified-site	/note= "Casein kinase II phosphorylation site"
FT	Modified-site	129. .153
FT	Modified-site	/note= "Protein kinase ATP-binding region signature"

26	1417	58.8	499	6	ABU61993	Human clk
27	1417	58.8	499	8	ADQ19736	Human sof
28	1417	58.8	499	8	ADQ15174	Human can
29	1400.5	58.1	352	5	AAE20171	Human kin
30	1331.5	55.2	497	2	AAW49912	Mouse CLK
31	1322	54.9	638	7	ADB64472	Human pro
32	1321	54.8	490	8	ADL12585	Human sec
33	1314.5	54.5	490	4	ABG23354	Novel hum
34	1310.5	54.4	431	4	AAAB85506	Human pro
35	1310	54.4	490	8	ADO08028	Mouse pol
36	1308	54.3	490	8	ADO08025	Human pol
37	1252	52.0	341	8	ADO08024	Human pol
38	1246	51.7	417	7	ADM06014	Human pro
39	1223.5	50.8	485	2	AAW49913	Mouse CLK
40	1222	50.7	511	4	ABB58352	Drosophil
41	1222	50.7	511	4	ABB58352	Drosophil
42	1222	50.7	511	8	ADO08023	Fly polyp
43	1051	43.6	265	6	ABR41275	Human DIT
44	848.5	35.2	425	3	AAG29704	Arabidops
45	848.5	35.2	427	3	AAG29703	Arabidops

XX New signal transducer and activator of transcription 6 activation
PT promoting purified protein, for diagnosing and treating disease
PT associated with activation/inhibition of transcription factor e.g.
PT diabetes and cancer.
XX
PS Claim 1; SEQ ID NO 108; 1368pp; English.
XX
The invention relates to a purified protein promoting signal transducer
and activator of transcription 6 activation (STAT6). The protein is
useful for the producing an antibody, which involves administering the
protein or its epitope-bearing fragments to a non-human animal as an
antigen. The nucleic acid is useful for diagnosing a disease or
susceptibility to a disease related to expression or activity of the
protein. A transformant expressing the protein is useful for screening
compounds which inhibit or promote STAT6 activation. A transformant
expressing the protein is useful for producing a pharmaceutical
composition. Compositions, antibodies and antisense molecules are useful
for the treating a disease associated with STAT6 activation such as
allergic diseases, inflammation, autoimmune diseases, diabetes,
hyperlipidaemia, infections disease and cancers. Compositions are useful
for treating disease associated with STAT6 activation and/or prevention
of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
viral hepatitis and AIDS. The protein has efficient promoting STAT6
activity. The protein or nucleic acid is effectively useful for screening
compounds for treating and preventing disease associated with excessive
activation or inhibition of STAT6. The present sequence represents the
amino acid sequence of a human protein which promotes STAT6 activation.

XX Sequence 481 AA;

Query Match 96.1%; Score 2315.5; DB 8; Length 481;
Best Local Similarity 96.2%; Pred. No. 3.5e-212;
Matches 432; Conservative 2; Mismatches 4; Indels 11; Gaps 2;
Qy 8 SH-SVEEDTH-----PSHYLEARSLSNERDYDRRYVDEYRNDYCEGVVPRHYRD 56
Db 33 SHSSTQENRHCKPHQPKESDCHYLEARSLSNERDYDRRYVDEYRNDYCEGVVPRHYRD 92
Qy 57 IESGVRHICSSSVRSRSPKRNHCHSSHQSKSHRRKRSIEDDEGHLCQSG 116
Db 93 IESGVRHICSSSVRSRSPKRNHCHSSHQSKSHRRKRSIEDDEGHLCQSG 152
Qy 117 DVLRYVEIVDTLGEAGFKVCEIDHGMGDMHVAVKIKVKNVGRYREARSEIQVLEHLN 176
Db 153 DVLRYVEIVDTLGEAGFKVCEIDHGMGDMHVAVKIKVKNVGRYREARSEIQVLEHLN 212
Qy 177 STDPSNVCQMLEWFDDHGHVCTVFFELGLSTYDFIKENSFLPFQIDHIRQWAYQICQ 236
Db 213 STDPSNVCQMLEWFDDHGHVCTVFFELGLSTYDFIKENSFLPFQIDHIRQWAYQICQ 272
Qy 237 SINFLHNKLTHTDLKPNILFKSDYVYKNSKMKRDLTKNTDIKVPDFGSATYDDE 296
Db 273 SINFLHNKLTHTDLKPNILFKSDYVYKNSKMKRDLTKNTDIKVPDFGSATYDDE 332
Qy 297 HHSTLVSTRHRAPEVILALGNSQPCDWSGTCILIEYLGFTVFTQTHDSKEHLAMMERI 356
Db 333 HHSTLVSTRHRAPEVILALGNSQPCDWSGTCILIEYLGFTVFTQTHDSKEHLAMMERI 392
Qy 357 LGPIPOHMIQTRKKYFHHNQLDWEHSSAGRYVRRCKPKLKEPMLCHDEHEKLPDLV 416
Db 393 LGPIPOHMIQTRKKYFHHNQLDWEHSSAGRYVRRCKPKLKEPMLCHDEHEKLPDLV 452
Qy 417 RRMLEYDPTQRTITLDEALQHPFFDLKKK 445
Db 453 RRMLEYDPTQRTITLDEALQHPFFDLKKK 481

RESULT 3
ADO08026
ID ADO08026 standard; protein; 481 AA.

XX ADO08026;
XX 01-JUL-2004 (first entry)
XX Human polypeptide #39.
XX Human; fat cell number; fat cell size; obesity; diabetes; anorectic;
XX antidiabetic.
XX Homo sapiens.
XX US2004071700-A1.
XX 15-APR-2004.
XX 09-OCT-2002; 2002US-00267502.
XX 09-OCT-2002; 2002US-00267502.
XX (LIFE-) LIFE SCI DEV CORP.
XX Kim J, Galant R;
XX WPI; 2004-328526/30.
XX N-PSDB; ADO07809.
XX Identifying compounds that influence fat cell number or size for treating
or preventing obesity or diabetes by exposing the cell to the agent and
identifying fat cell number or size relative to cells not exposed to the
agent.
XX
XX Claim 14; SEQ ID NO 352; 275pp; English.
XX
The invention relates to a method of identifying compounds that influence
fat cell number or size comprising providing a cell that expresses a gene
and an agent, exposing the cell to the agent and identifying fat cell
number or size relative to cells not exposed to the agent. The method
also comprises providing an expression vector and an agent, exposing the
vector to the agent, detecting a change in expression of the gene
relative to expression of the gene in an expression vector not exposed to
the agent, treating a subject with the agent and identifying fat cell
number or size in the subject. The agent comprises an antisense
oligonucleotide. The subject comprises a mammal, preferably a human. The
method also comprises providing a polypeptide and an agent, exposing the
polypeptide to the agent, detecting binding of the agent to the
polypeptide or a change in an activity of the polypeptide, treating a
subject with the agent and identifying fat cell number or size in the
subject. The agent comprises an antibody. A method of regulating fat cell
number or size comprises providing a subject containing fat cells and an
agent that changes the expression of a gene, and treating the subject
with the agent under conditions so that fat cell size or number in the
subject is altered. The method is useful for identifying compounds that
influence fat cell number or size, for preparing a composition for
treating or preventing obesity or diabetes. This sequence represents a
human polypeptide used in the scope of the invention.

XX Sequence 481 AA;

Query Match 96.1%; Score 2315.5; DB 8; Length 481;
Best Local Similarity 96.2%; Pred. No. 3.5e-212;
Matches 432; Conservative 2; Mismatches 4; Indels 11; Gaps 2;
Qy 8 SH-SVEEDTH-----PSHYLEARSLSNERDYDRRYVDEYRNDYCEGVVPRHYRD 56
Db 33 SHSSTQENRHCKPHQPKESDCHYLEARSLSNERDYDRRYVDEYRNDYCEGVVPRHYRD 92
Qy 57 IESGVRHICSSSVRSRSPKRNHCHSSHQSKSHRRKRSIEDDEGHLCQSG 116
Db 93 IESGVRHICSSSVRSRSPKRNHCHSSHQSKSHRRKRSIEDDEGHLCQSG 152
Qy 117 DVLRYVEIVDTLGEAGFKVCEIDHGMGDMHVAVKIKVKNVGRYREARSEIQVLEHLN 176

Db 153 DVLARVEIVDTLGEAGFKVVEICDHGMGMHVAVKIVKNGVRYREARSEIQVLEHLN 212
 Qy 177 STDPSVRCVQMLEWFDHGHVCIIVFELLGLSYDFIKENSFLPFOIDHIROMAYOICQ 236
 Db 213 STDPSVRCVQMLEWFDHGHVCIIVFELLGLSYDFIKENSFLPFOIDHIROMAYOICQ 272
 Qy 237 SINFLHNKLTHDLPENILFVKSDDYVYVYKNSKMKRDERTLKNVDIKNVDFGSAFYDDE 296
 Db 273 SINFLHNKLTHDLPENILFVKSDDYVYVYKNSKMKRDERTLKNVDIKNVDFGSAFYDDE 332
 Qy 297 HSTLVSTRHYRAPEVILALGWSQPCDVMSIGCIIIEYILGFTVFOHDSKEHLAMMERI 356
 Db 333 HSTLVSTRHYRAPEVILALGWSQPCDVMSIGCIIIEYILGFTVFOHDSKEHLAMMERI 392
 Qy 357 LGPIPOHMIQTRKRYFHHNQLDWEHSSAGRYVRRCKPLKEFMLCHDEEHEKLFDLV 416
 Db 393 LGPIPOHMIQTRKRYFHHNQLDWEHSSAGRYVRRCKPLKEFMLCHDEEHEKLFDLV 452
 Qy 417 RRMLEYDPTORITLDEALQHPFFDLKKK 445
 Db 453 RRMLEYDPTORITLDEALQHPFFDLKKK 481

RESULT 4
 ADP24394
 ID ADP24394 standard; protein; 481 AA.
 AC ADP24394;
 XX
 DT 18-NOV-2004 (first entry)
 DE PRO polypeptide SEQ ID NO:1572.
 KW PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
 KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
 KW antiaschmatic; hepatotropic; respiratory; gene therapy; immune system.
 OS Unidentified.
 XX WO2004041170-A2.
 XX 21-MAY-2004.
 XX 30-OCT-2003; 2003WO-US034312.
 XX 01-NOV-2002; 2002US-0423394P.
 XX (GETH) GENENTECH INC.
 XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WT;
 PI Wu TD;
 XX WPI; 2004-419628/39.
 XX N-PSDB; ADP24393.
 XX

New PRO polypeptides and polynucleotides, useful for treating e.g.
 PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
 PT renal disease, or demyelinating diseases of the central or peripheral
 PT nervous system.
 XX
 PS Claim 7; SEQ ID NO 1572; 2940pp; English.
 XX
 CC The invention relates to a novel isolated nucleic acid and the PRO
 CC polypeptide encoded by it. A protein of the invention has
 CC antiinflammatory, antirheumatic, antipsoriatic, immunosuppressive,
 CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
 CC antiaschmatic, hepatotropic, and respiratory activity. A polynucleotide
 CC of the invention may have a use in gene therapy. The PRO polypeptide, its
 CC agonist, antagonist, or antibody that specifically binds to the
 CC polypeptide is useful for treating an immune related disorder such as
 CC systemic lupus erythematous, rheumatoid arthritis, osteoarthritis,
 CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
 CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic

CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
 CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
 CC disease, a demyelinating disease of the central or peripheral nervous
 CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
 CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
 CC disease, infectious or autoimmune chronic active hepatitis, primary
 CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
 CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
 CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
 CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
 CC disease, asthma, allergic rhinitis, atopic dermatitis, food
 CC hypersensitivity, urticaria, an immunologic disease of the lung,
 CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
 CC pneumonitis, a transplantation associated disease, graft rejection or
 CC graft-versus-host disease. The present sequence represents a PRO protein
 CC of the invention.
 XX
 SQ Sequence 481 AA;
 Query Match 96.1%; Score 2315.5; DB 8; Length 481;
 Best Local Similarity 96.2%; Pred. No. 3.5e-212;
 Matches 432; Conservative 2; Mismatches 4; Indels 11; Gaps 2;
 Qy 8 SH-SVERDTH-----PSHYLEARSINERDYDRRYVDEYRNDYCEGYVPRHYRD 56
 Db 33 SHSSTQENRCHKPHQFQKESDCHYLEARSINERDYDRRYVDEYRNDYCEGYVPRHYRD 92
 Qy 57 IESGYRIHCSKSVRRSRSPKRNHCSHOSKSRSHRRKRSRSTIEDDEGHLCQSG 116
 Db 93 IESGYRIHCSKSVRRSRSPKRNHCSHOSKSRSHRRKRSRSTIEDDEGHLCQSG 152
 Qy 117 DVLARVEIVDTLGEAGFKVVEICDHGMGMHVAVKIVKNGVRYREARSEIQVLEHLN 176
 Db 153 DVLARVEIVDTLGEAGFKVVEICDHGMGMHVAVKIVKNGVRYREARSEIQVLEHLN 212
 Qy 177 STDPSVRCVQMLEWFDHGHVCIIVFELLGLSYDFIKENSFLPFOIDHIROMAYOICQ 236
 Db 213 STDPSVRCVQMLEWFDHGHVCIIVFELLGLSYDFIKENSFLPFOIDHIROMAYOICQ 272
 Qy 237 SINFLHNKLTHDLPENILFVKSDDYVYVYKNSKMKRDERTLKNVDIKNVDFGSAFYDDE 296
 Db 273 SINFLHNKLTHDLPENILFVKSDDYVYVYKNSKMKRDERTLKNVDIKNVDFGSAFYDDE 332
 Qy 297 HSTLVSTRHYRAPEVILALGWSQPCDVMSIGCIIIEYILGFTVFOHDSKEHLAMMERI 356
 Db 333 HSTLVSTRHYRAPEVILALGWSQPCDVMSIGCIIIEYILGFTVFOHDSKEHLAMMERI 392
 Qy 357 LGPIPOHMIQTRKRYFHHNQLDWEHSSAGRYVRRCKPLKEFMLCHDEEHEKLFDLV 416
 Db 393 LGPIPOHMIQTRKRYFHHNQLDWEHSSAGRYVRRCKPLKEFMLCHDEEHEKLFDLV 452
 Qy 417 RRMLEYDPTORITLDEALQHPFFDLKKK 445
 Db 453 RRMLEYDPTORITLDEALQHPFFDLKKK 481
 RESULT 5
 AAB69705
 ID AAB69705 standard; protein; 481 AA.
 XX AAB69705;
 XX
 DT 14-AUG-2001 (first entry)
 DE Human cell cycle regulating protein 53.
 XX
 KW Human; cell cycle regulating protein 53; cancer; blood disease; HIV;
 KW immunological disease; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200130833-A1.
 XX

93	Db	VESTYRIHCSKSVSRSRSSPKRKNRPCCASHQSHSKSHRRKGRSRSTEDDEGHLLCQSG	152
117	Qy	DVLRYEYETDVLGEGAKGVVECIDHGMQGMHVAVIKVNGRYREAAARSEIQVLEHLN	176
153	Db	DVLRYEYETDVLGEGAKGVVECIDHGMQGMHLVAVIKVNGRYREAAARSEIQVLEHLN	212
177	Qy	STDPNSVFCVQWLEWFODHHGVCIVFPELLGLSTYDFIKENSFPLPFOIDHIRMAYQICQ	236
213	Db	STDPNSVFCVQWLEWFODHHGVCIVFPELLGLSTYDFIKENSFPLPFOIDHIRMAYQICQ	272
237	Qy	SINFLHNKLTHTDLKPENILFVKSDDYVVKNSNMKDERLTKNXTDKVVDGFSATYDDE	296
273	Db	SINFLHNKLTHTDLKPENILFVKSDDYVVKNSNMKDERLTKNXTDKVVDGFSATYDDE	332
297	Qy	HSTLSTRHYRAPEVILALGWSQPCDWSIGCILIEYILGFTVFTQTHDSKEHLAMWERI	356
333	Db	HSTLSTRHYRAPEVILALGWSQPCDWSIGCILIEYILGFTVFTQTHDSKEHLAMWERI	392
357	Qy	LGP I PQHMIQTKRKRYFHHNQLDWDEHSSAGRYVRRRCCKPLKEFMLCHDDEHEKLFDLV	416
393	Db	LGP I PAHMIQTKRKRYFHHNQLDWDEHSSAGRYVRRRCCKPLKEFMLCHDDEHEKLFDLV	452
417	Qy	RMLEVDPTQRTILDEALQHPFFDLLKKK	445
453	Db	RMLEVDTPARRITTLDEALQHPFFDLLKKK	481

```

RESULT 7
ADO08029
ID ADO08029 standard; protein; 481 AA.
XX
XX ADO08029;
XX
XX 01-JUL-2004 (first entry)
DT
XX
XX Mouse polypeptide #36.
XX
XX Mouse; fat cell number; fat cell size; obesity; diabetes; anorectic;
KW antidiabetic.
KW
XX
XX Mus sp.
OS
XX
XX US2004071700-A1.
XX
XX 15-APR-2004.
XX
XX
XX 03-OCT-2002; 2002US-00267502.
PF
XX
XX 03-OCT-2002; 2002US-00267502.
XX
XX (LIFE-) LIFE SCI DEV CORP.
XX
XX
XX Kim J, Galant R;
PI
XX
XX WPI; 2004-328526/30.
DR N-PSDS; ADO07812.
DR

```

number or size in the subject. The agent comprises an antisense oligonucleotide. The subject comprises a mammal, preferably a human. The method also comprises providing a polypeptide and an agent, exposing the polypeptide to the agent, detecting binding of the agent to the polypeptide or a change in an activity of the polypeptide, treating a subject with the agent and identifying fat cell number or size in the subject. The agent comprises an antibody. A method of regulating fat cell number or size comprises providing a subject containing fat cells and an agent that changes the expression of a gene, and treating the subject with the agent under conditions so that fat cell size or number in the subject is altered. The method is useful for identifying compounds that influence fat cell number or size, for preparing a composition for treating or preventing obesity or diabetes. This sequence represents a mouse polypeptide used in the scope of the invention.

RESULT	8
AAM49914	
ID	AAW49914 standard; protein; 479 AA.
XX	
XX	AAW49914;
XX	
XX	
DT	20-JUL-1998 (first entry)
XX	
DE	Mouse CLK serine/threonine kinase mCLK4.
XX	
XX	mCLK4; CLK; serine/threonine kinase; protein kinase; LAMMER kinase; signal transduction; cancer; contraceptive; mouse; therapy; diagnosis
KW	
XX	
OS	Mus musculus.
XX	
Key	Location/Qualifiers
FH	29..44
Domain	/note= "nuclear localisation domain"
FT	159..476
Domain	
FT	

FT	Peptide	/note= "catalytic domain"
FT	PT	384...389
FT	PT	/note= "LAMMER motif"
PN	XX	WO9748723-A2...
XX	XX	24-DEC-1997.
PD	XX	
XX	XX	17-JUN-1997; 97WO-IB000946.
XX	XX	
PR	17-JUN-1996;	96US-0019629P.
PR	09-AUG-1996;	96US-0023485P.
PR	13-NOV-1996;	96US-0030860P.
PR	15-NOV-1996;	96US-0030964P.
PR	19-DEC-1996;	96US-0034286P.
XX	XX	
PA	(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.	
PI	Ullrich A, Kharitonov AI, Aoki N, Wang HY, Chen Z, Nayler O;	
PI	Klim YW;	
XX	XX	WPI; 1998-120302/11.
XX	XX	
PT	New phosphatase and kinase enzyme(s) - useful in the diagnosis and	
PT	treatment of signal transduction disorders.	
XX	XX	
PS	Claim 11; Fig 4; 138pp; English.	
XX	XX	
CC	This polypeptide comprises novel mouse CLK serine/threonine kinase mCLK4,	
CC	from the CLK serine/threonine kinase family of proteins that regulate RNA	
CC	splicing in cells. mCLK4 cDNA was cloned from a mouse embryo 11.5 p.c.	
CC	12AP cDNA library. The invention discloses the discovery of novel protein	
CC	kinases mCLK2 (see AAW49912), mCLK3 and mCLK4 (see AAW49914) of mol.wt.	
CC	59.9, 58.5 and 57.2 kDa, respectively, as well as other novel proteins	
CC	(see AAW49906-10) involved in cellular signal transduction, and provides	
CC	vectors, host cells, purified recombinant proteins, methods for	
CC	identifying compounds that activate or inhibit the novel proteins, as	
CC	well as methods for the diagnosis and treatment of diseases associated	
CC	with the novel proteins. Overexpression of CLK serine/threonine kinases	
CC	has been implicated in certain types of cancer. Compounds that inhibit	
CC	their catalytic activity or disrupt their interactions with natural	
CC	binding partners may act as anti-cancer therapeutics. mCLK related	
CC	molecules and compounds may also be useful as male contraceptives	
XX	XX	
SQ	Sequence 479 AA;	
Query Match	90.4%; Score 2177.5; DB 2; Length 479;	
Best Local Similarity	91.1%; Pred. No. 5.4e-199;	
Matches 409; Conservative	9; Mismatches 18; Indels 13; Gaps 4	
QY	8 SH-SVEEDTH-----PSHYLEARSLSNERDYDRRYVDEYRNDYCEGVPRHYRD 56	
Db	33 SHSTQENRCKPHQPKDSCHVLEARCLNERDYRD-RYDEYRNDYCEGVPRHYRD 91	
QY	57 IEGYRIHCSKSSVRRSSRPKRKNRHCSHQSKSRKRSRSIEDDEEGHLICQSG 116	
Db	92 VESTYRIHCSKSSVRRSSRPKRKNRPFASHQSHSHRRKRKRSRSIEDDEEGHLICQSG 151	
QY	117 DVLRLARYEIVDTLGEAGFGKVEICIDHGMQGHVAVKIVKNGRYREARSEIQVLEHLN 176	
Db	152 DVLRLARYEIVDTLGEAGFGKVEICIDHGMQGLHVAVKIVKNGGYREARSEIQVLEHLN 211	
QY	177 STDNSVFRVCQMLEWFDHGHGVCIVFELLGLISTYDFIKENSFLPFQIDHIRMAYQICQ 236	
Db	212 STDNSVFRVCQMLEWFDHGHGVCIVFELLGLISTYDFIKENSFLPFQIDHIRMAYQICQ 271	
QY	237 SINFLHKNLTHDLKPENILFVKSDDYVVKYKNSMKRDERLTAKNTDIKVPDFGSATYDDE 296	
Db	272 SINFLHKNLTHDLKPENILFVKSDDYVVK-NPSMKRDERILTPTIKVDFGSATYDDE 330	
QY	297 HHSTLVSTRHYRAPEVIALGWSQPCDWSIGCILIEYLLGFTVPTQTHDSKEHLAMMERI 356	
Db	331 HHSTLVSTRHYRAPEVIALGWSQPCDWSIGCILIEYLLGFTVPTQTHDSKEHLAMMERI 390	

QY 8 SH-SVEEDTH-----PSHYLEARSINERDYYVDEYRNDYCEGVVPRHYRD 56
 DB 81 SHSSTQENRHCKPHQPKESDCHYLEARSINERDYYVDEYRNDYCEGVVPRHYRD 140
 QY 57 IESGYRIHCSKSSVRSRRSPKRNHCHSHQSRSHRRKRSRSDDEEGHLICQSG 116
 DB 141 IESGYRIHCSKSSVRSRRSPKRNHCHSHQSR----- 176
 QY 117 DVLBARVEIVDTLGEAGFGKVEICIDHGMGMHVAVKIVKNGRYREARSEIOVLEHLN 176
 DB 177 -----XEIVDTLGEAGFGKVEICIDHGMGMHVAVKIVKNGRYREARSEIOVLEHLN 230
 QY 177 STDNSVFRVCQMLEWFDHGHGVCIVFELLGLSTYDFIKENSFLPQIDHIRQWAYQICQ 236
 DB 231 STDNSVFRVCQMLEWFDHGHGVCIVFELLGLSTYDFIKENSFLPQIDHIRQWAYQICQ 290
 QY 237 SINFLHNKLTHTDLKPENILFKVSDYVVKYNSKMKRDETLKNTDIKVDFGSGATYDDE 296
 DB 291 SINFLHNKLTHTDLKPENILFKVSDYVVKYNSKMKRDETLKNTDIKVDFGSGATYDDE 350
 QY 297 HHSTLVSTRHYRAPEVILALGWSQPCDVWSIGCILIEYILGFTVFTQTHDSKEHLAMMERI 356
 DB 351 HHSTLVSTRHYRAPEVILALGWSQPCDVWSIGCILIEYILGFTVFTQTHDSKEHLAMMERI 410
 QY 357 LGPIPOHMIQKTRKRYFHHNQLDWEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLPDLV 416
 DB 411 LGPIPOHMIQKTRKRYFHHNQLDWEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLPDLV 470
 QY 417 RRMLEYDPTQRTITLDEALQHPFFDLKKK 445
 DB 471 RRMLEYDPTQRTITLDEALQHPFFDLKKK 499

RESULT 10
 ADI29255
 ID ADI29255 standard; protein; 499 AA.
 AC ADI29255;
 XX
 DT 22-APR-2004 (first entry)
 DE Human MARK3-associated protein #40.
 XX
 KW Human; antisense gene therapy; MARK3;
 KW MAP/microtubule affinity-regulating kinase 3; cancer;
 KW Alzheimer's disease; neurodegenerative disorder;
 KW hyperproliferative disorder; cytostatic.
 OS Homo sapiens.
 XX
 PN US2003232771-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 17-JUN-2002; 2002US-00174319.
 XX
 PR 17-JUN-2002; 2002US-00174319.
 XX
 PA (ISIS-) ISIS PHARM INC.
 XX
 PI Ward DT, Freier SM, Dobie KW;
 XX
 DR WPI; 2004-052188/05.
 DR N-FSDB; ADI29373.
 XX
 XX New antisense compound targeted to a nucleic acid molecule encoding
 PT microtubule-affinity-regulating kinases (MARK3), useful for modulating
 PT expression of MARK3 or for treating cancer or Alzheimer's disease.
 XX
 PS Disclosure; SEQ ID NO 175; 233pp; English.
 XX
 CC The invention relates to a compound comprising a sequence comprising 8-80

CC base pairs (bp) targeted to a nucleic acid encoding MARK3
 CC (MAP/microtubule affinity-regulating kinase 3), that specifically
 CC hybridizes with the nucleic acid encoding MARK3 and inhibits expression
 CC of MARK3, i.e. is an antisense oligonucleotide (AO). Also included are a
 CC composition comprising the compound and a carrier or diluent, inhibiting or
 CC the expression of MARK3 in cells or tissues, treating an animal having or
 CC suspected of having a disease or condition associated with MARK3 and
 CC screening for an antisense compound. The antisense oligonucleotide is
 CC useful for preparing a composition for treating hyperproliferative
 CC disorder, particularly cancer and neurodegenerative diseases e.g.
 CC Alzheimer's disease. The present sequence is a MARK3 associated protein
 CC included in the figures but not mentioned anywhere else in the
 CC specification.
 XX Sequence 499 AA;
 SQ

Query Match 88.3%; Score 2127.5; DB 8; Length 499;
 Best Local Similarity 89.3%; Pred. No. 3.5e-194;
 Matches 401; Conservative 2; Mismatches 5; Indels 41; Gaps 3;

QY 8 SH-SVEEDTH-----PSHYLEARSINERDYYVDEYRNDYCEGVVPRHYRD 56
 DB 81 SHSSTQENRHCKPHQPKESDCHYLEARSINERDYYVDEYRNDYCEGVVPRHYRD 140
 QY 57 IESGYRIHCSKSSVRSRRSPKRNHCHSHQSRSHRRKRSRSDDEEGHLICQSG 116
 DB 141 IESGYRIHCSKSSVRSRRSPKRNHCHSHQSR----- 176
 QY 117 DVLBARVEIVDTLGEAGFGKVEICIDHGMGMHVAVKIVKNGRYREARSEIOVLEHLN 176
 DB 177 -----XEIVDTLGEAGFGKVEICIDHGMGMHVAVKIVKNGRYREARSEIOVLEHLN 230
 QY 177 STDNSVFRVCQMLEWFDHGHGVCIVFELLGLSTYDFIKENSFLPQIDHIRQWAYQICQ 236
 DB 231 STDNSVFRVCQMLEWFDHGHGVCIVFELLGLSTYDFIKENSFLPQIDHIRQWAYQICQ 290
 QY 237 SINFLHNKLTHTDLKPENILFKVSDYVVKYNSKMKRDETLKNTDIKVDFGSGATYDDE 296
 DB 291 SINFLHNKLTHTDLKPENILFKVSDYVVKYNSKMKRDETLKNTDIKVDFGSGATYDDE 350
 QY 297 HHSTLVSTRHYRAPEVILALGWSQPCDVWSIGCILIEYILGFTVFTQTHDSKEHLAMMERI 356
 DB 351 HHSTLVSTRHYRAPEVILALGWSQPCDVWSIGCILIEYILGFTVFTQTHDSKEHLAMMERI 410
 QY 357 LGPIPOHMIQKTRKRYFHHNQLDWEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLPDLV 416
 DB 411 LGPIPOHMIQKTRKRYFHHNQLDWEHSSAGRYVRRRCCKPLKEFMLCHDEHEKLPDLV 470
 QY 417 RRMLEYDPTQRTITLDEALQHPFFDLKKK 445
 DB 471 RRMLEYDPTQRTITLDEALQHPFFDLKKK 499

RESULT 11
 AAY27054
 ID AAY27054 standard; protein; 451 AA.
 XX
 AC AAY27054;
 XX
 DT 08-OCT-1999 (first entry)
 DE Human protein kinase (HPKM)-3 (clone ID 339963).
 XX
 DE Human protein kinase molecule; HPKM; human; protein kinase;
 KW phosphate group; cancer; immune disorder.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 4
 FT Modified-site 22 /note= "potential phosphorylation site"
 FT Modified-site 22 /note= "potential phosphorylation site"

FT Modified-site 26 /note= "potential phosphorylation site"
FT Modified-site 33 /note= "potential phosphorylation site"
FT Modified-site 36 /note= "potential phosphorylation site"
FT Modified-site 56 /note= "potential phosphorylation site"
FT Modified-site 61 /note= "potential phosphorylation site"
FT Modified-site 67 /note= "potential phosphorylation site"
FT Modified-site 105 /note= "potential phosphorylation site"
FT Modified-site 108 /note= "potential phosphorylation site"
FT Modified-site 112 /note= "potential phosphorylation site"
FT Modified-site 128 /note= "potential phosphorylation site"
FT Modified-site 134 /note= "potential phosphorylation site"
FT Domain 135..159 /note= "potential phosphorylation site"
FT /note= "potential catalytic domain"
FT Modified-site 182 /note= "potential N-glycosylation site"
FT Modified-site 215 /note= "potential phosphorylation site"
FT Domain 252..264 /note= "potential catalytic domain"
FT Modified-site 253 /note= "potential phosphorylation site"
FT Modified-site 283 /note= "potential phosphorylation site"
FT Modified-site 298 /note= "potential phosphorylation site"
FT Modified-site 309 /note= "potential phosphorylation site"
FT Modified-site 374 /note= "potential phosphorylation site"
FT Modified-site 431 /note= "potential phosphorylation site"
FT Modified-site 435 /note= "potential phosphorylation site"
XX WO9938981-A2.
XX
XX 05-AUG-1999.
XX 12-JAN-1999; 99WO-US000661.
XX 30-JAN-1998; 98US-00016000.
XX (INCY-) INCYTE PHARM INC.
XX Bandman O, Hillman JL, Lal P, Akerblom IE, Shah P, Corley NC;
XX Guegler KJ;
XX
XX WPI; 1999-479190/40.
XX N-PSDB; AAX89852.
XX
XX New human protein kinase molecules useful for treating or preventing
XX cancer or an immune disorder.
XX Claim 1; Page 67-68; 77pp; English.
XX The invention provides human protein kinase molecules (HPKM) (AAY27052-
XX 57) and nucleic acid sequences (AAX89850-55) encoding the HPKM
XX polypeptides respectively. The HPKM polypeptides can be produced
XX recombinantly by standard recombinant methodology. Protein kinases add
XX phosphate groups to proteins. HPKM polypeptide or an HPKM antagonist is

CC used to treat or prevent cancer or an immune disorder. The present
CC sequence represents the amino acid sequence of HPKM-3
XX
SQ Sequence 451 AA;
Query Match 87.9%; Score 2117.5; DB 2; Length 451;
Best Local Similarity 88.9%; Pred. No. 2.7e-193;
Matches 399; Conservative 3; Mismatches 6; Indels 41; Gaps 3;
Qy 8 SH-SVEEDTH-----PSHYLEARSLSNERDYDRRYVDEYRNDYCEGYVPRHYRD 56
||| : |||
Db 33 SHSSTQENRCKPHQPKESDCHYLEARSLSNERDYDRRYVDEYRNDYCEGYVPRHYRD 92
||| : |||
Qy 57 IESGYRIHCSKSSVRSRRSPKRNHRHCSHQSRKSRKRSIEDDEBEGHLICSG 116
||| : |||
Db 93 IESGYRIHCSKSSVRSRRSPKRNHRHCSHQSRKSRKRSIEDDEBEGHLICSG 128
||| : |||
Qy 117 DVLRARYEIVDTLGEAGFGKVECIDHGMGMHVAVKXNVGRYREARSEIOVLEHLN 176
||| : |||
Db 129 -----MKSVDTLGEAGFGKVECIDHGMGMHVAVKXNVGRYREARSEIOVLEHLN 182
||| : |||
Qy 177 STDNSVRCVQMLEWFDHGHVCIVFELLGLSTYDFIKENSFLPFQIDHIRMAYQICQ 236
||| : |||
Db 183 STDNSVRCVQMLEWFDHGHVCIVFELLGLSTYDFIKENSFLPFQIDHIRMAYQICQ 242
||| : |||
Qy 237 SINFLHNKLTHTDLKPNILFVKSDYVVKYNSQMKRDLTKNTDIKVDFGSGATYDDE 296
||| : |||
Db 243 SINFLHNKLTHTDLKPNILFVKSDYVVKYNSQMKRDLTKNTDIKVDFGSGATYDDE 302
||| : |||
Qy 297 HSTLSTVSTRYRADEVILALGWSQPCDVWSIGCILIBYILGFTVFTQTHDSKEHLAMMERI 356
||| : |||
Db 303 HSTLSTVSTRYRADEVILALGWSQPCDVWSIGCILIBYILGFTVFTQTHDSKEHLAMMERI 362
||| : |||
Qy 357 LGPIPOHMIQKTRKRYFHNQLDWDEHSSAGRYVRRRCRCKPLKEFMLCHDEHEKLPDLV 416
||| : |||
Db 363 LGPIPOHMIQKTRKRYFHNQLDWDEHSSAGRYVRRRCRCKPLKEFMLCHDEHEKLPDLV 422
||| : |||
Qy 417 RRMLEYDPTQRTILDEALQHPFFDLKKK 445
||| : |||
Db 423 RRMLEYDPTQRTILDEALQHPFFDLKKK 451
RESULT 12
ADO08027
ID ADO08027 standard; protein; 484 AA.
XX
XX ADO08027;
XX
XX 01-JUL-2004 (first entry)
XX Human polypeptide #40.
XX
XX Human; fat cell number; fat cell size; obesity; diabetes; anorectic;
XX antidiabetic.
XX
XX Homo sapiens.
XX
XX US2004071700-A1.
XX
XX 15-APR-2004.
XX
XX 09-OCT-2002; 2002US-00267502.
XX
XX 09-OCT-2002; 2002US-00267502.
XX
XX (LIFE-) LIFE SCI DEV CORP.
XX
XX Kim J, Galant R;
XX
XX WPI; 2004-328526/30.
XX N-PSDB; ADO07810.
XX
XX Identifying compounds that influence fat cell number or size for treating

or preventing obesity or diabetes by exposing the cell to the agent and identifying fat cell number or size relative to cells not exposed to the agent.

Claim 14; SEQ ID NO 353; 275pp; English.

The invention relates to a method of identifying compounds that influence fat cell number or size comprising providing a cell that expresses a gene and an agent, exposing the cell to the agent and identifying fat cell number or size relative to cells not exposed to the agent. The method also comprises providing an expression vector and an agent, exposing the vector to the agent, detecting a change in expression of the gene relative to expression of the gene in an expression vector not exposed to the agent, treating a subject with the agent and identifying fat cell number or size in the subject. The agent comprises an antisense oligonucleotide. The subject comprises a mammal, preferably a human. The method also comprises providing a polypeptide and an agent, exposing the polypeptide to the agent, detecting binding of the agent to the polypeptide or a change in an activity of the polypeptide, treating a subject with the agent and identifying fat cell number or size in the subject. The agent comprises an antibody. A method of regulating fat cell number or size comprises providing a subject containing fat cells and an agent that changes the expression of a gene, and treating the subject with the agent under conditions so that fat cell size or number in the subject is altered. The method is useful for identifying compounds that influence fat cell number or size, for preparing a composition for treating or preventing obesity or diabetes. This sequence represents a human polypeptide used in the scope of the invention.

Sequence 484 AA;

Query Match 78.5%; Score 1893; DB 8; Length 484;
 Best Local Similarity 82.3%; Pred. No. 8.9e-172;
 Matches 353; Conservative 30; Mismatches 44; Indels 2; Gaps 2;
 18 SHYLEARSLNDRYDRDYDEYNDYCEGYVPRHYRDIESGYRHCSSKSVSRSSSP 77
 54 SHYLESRSINEKDVHSRRYIDYRNDYTGCGEPGHRQDHESRYQNHSSKSGSRSSY 113
 78 KKK-RNRCHSH-QSRKSHRKRKRSRSTEDDEEGLHCQSGDVLRAVEIVDTLGEAGF 135
 114 KSKHRIHSTSHRSRSHGSKSRKRKTRSVEDDEEGLHCQSGDVLRAVEIVDTLGEAGF 173
 136 KVEICIDHGMGMVAVKIVKNGRYREARSEIQVLEHLNSTDPNSVFCVQMLEWFDH 195
 174 KVVSCIDHKAGRHVAVKIVKNDVRYCEARSEIQVLEHLNTDPNSTFCVQMLEWFEH 233
 196 HGHVCIPELLGLSTYDFIKENSLPFDQIDHIRMAYQICQINFLHNNKLTHTDLPEN 255
 234 HGHICIVPELLGLSTYDFIKENGLPFLDHRKMAQYQICKSVNFLSHNKLTHTDLPEN 293
 256 ILFVKSDDYVKNYSKMKRDETLKNTDKVDFGSATYDDDEHHTVSTRHYRAPEVILA 315
 294 ILFVQSDYTEANPKIKRDETLNPKDKVDFGSATYDDDEHHTVSTRHYRAPEVILA 353
 316 LGWQPCDVMWIGICILLEYLGFVTFVTHDSKEHLAWERILGIPQHMIOKTRKRYFH 375
 354 LGWQPCDVMWIGICILLEYLGFVTFVTHDSKEHLAWERILGIPKHMIOKTRKRYFH 413
 376 HNQLDWEHSHSAGRYVRRCKPLKEFMLCHDEHEKFLDLVRRMLVEYDPTQRIFLDALQ 435
 414 HDRLDWEHSHSAGRYVRRCKPLKEFMLSQDVEHERLFDLIQKMLEYDPAKRITLREALK 473
 436 HPFFDLKK 444
 474 HPFFDLKK 482

RESULT 13
 AAU75108 standard; protein; 484 AA.
 XX AC AAU75108;

09-APR-2002 (first entry)
 Serine and Threonine protein kinase Clk.
 MAPKAP-K3; AP-3 delta; APP-695; amyloid A4; Hsp8; heat shock protein 8;
 LI30; NY-REN-58; P38 Alpha; P38 Beta; ERK3; KIAA0934; CDK9;
 cell cycle dependent kinase 9; CLK; C-NAP-1; clathrin heavy chain;
 amphiphysin; PN9109; KIAA1106; neurodegenerative disorder;
 non-insulin dependent diabetes mellitus; NIDDM; Alzheimer's disease;
 inflammatory disease; rheumatoid arthritis; inflammatory bowel disorder;
 atherosclerosis; cardiac hypertrophy; hypoxic brain injury;
 yeast two-hybrid; signal transduction pathway; human;
 mitogen activated protein kinase.
 Homo sapiens.
 OS
 Key Location/Qualifiers
 Region 1.364
 /note="This region binds the mitogen activated protein
 kinase (MAPK) ERK3 (see AAU75105), the bait protein in a
 yeast two-hybrid assay, producing a complex claimed in
 Claim 1"
 WO200198524-A2.
 27-DEC-2001.
 21-JUN-2001; 2001WO-US019762.
 22-JUN-2000; 2000US-0213245P.
 (MYRI-) MYRIAD GENETICS INC.
 Heichman K, Bartel PL;
 WPI; 2002-122287/16.
 New protein complexes comprising protein-protein interactions (e.g.
 MAPKAP-K3/AP-3 delta or C-NAP-1/Clathrin HC) useful for diagnosing
 physiological generative disorders or screening drugs for these diseases.
 Example 9; Page: 60pp; English.
 The invention describes an isolated protein complex, comprising two
 proteins. The protein complex comprises: protein kinase MAPKAP-K3 and AP-
 3 adaptor complex delta sub-unit; MAPKAP-K3 and amyloid A4 precursor
 protein, APP-695; MAPKAP-K3 and heat shock protein (Hsp) 8; leucine rich
 LI30 and NY-REN-58; P38 Alpha and P38 Beta; protein kinase ERK3 and
 KIAA0934 (unknown function); ERK3 and cell cycle dependent kinase (CDK) 9
 ; ERK3 and protein kinase CLK; C-NAP-1 and Clathrin heavy chain; C-NAP-1
 and Amphiphysin; C-NAP-1 and novel protein PN9109 or C-NAP-1 and KIAA1106
 (unknown function) interactions. The protein complexes are useful for
 diagnosing physiological generative disorders, drug screening for agents
 that modulate the interaction of the proteins (thus identify drug
 targets), and identifying additional proteins in the pathway common to
 the proteins. These physiological disorders include non-insulin dependent
 diabetes mellitus (NIDDM), neurodegenerative disorders (e.g. Alzheimer's
 disease), inflammatory diseases (e.g. rheumatoid arthritis and
 inflammatory bowel disorder) and other human disease such as
 atherosclerosis, cardiac hypertrophy and hypoxic brain injury. This
 sequence represents the serine and threonine kinase Clk, residues 1-364
 of which binds to the bait protein ERK3, a mitogen activated protein
 kinase (AAU75105) in a yeast two-hybrid assay for determining components
 of signal transduction pathways and forms an interaction claimed in claim
 1 of the invention. Note: This sequence does not appear in the
 CC specification but has been obtained from a reference given in the
 CC invention
 XX SQ
 Sequence 484 AA;
 Query Match 78.3%; Score 1887; DB 5; Length 484;
 Best Local Similarity 82.1%; Pred. No. 3.3e-171;

KW HIV propagation; human.
XX Homo sapiens.
XX WO2004065577-A2.
XX 05-AUG-2004.
XX 13-JAN-2004; 2004WO-US000798.
XX 14-JAN-2003; 2003US-0440068P.
XX 12-MAY-2003; 2003US-0469757P.
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX Nadler SG, Neubauer MG, Feder JN, Carman J;
XX WPI; 2004-562168/54.
XX N-ESDB; ADR14000.
XX New isolated polynucleotides and polypeptides associated with NF-kappaB
XX pathway, useful for diagnosing, treating, or preventing disorders or
XX diseases associated with NF-kappaB pathway.
XX Claim 6; SEQ ID NO 2; 237pp; English.
XX This invention relates to the novel association of protein sequences (and
XX the genes which encode them) to the NF-kappaB pathway. The invention may
XX be useful for the production of compounds with an anti-inflammatory,
XX cytostatic, hepatotropic, virucide, antiarthritic, antirheumatic,
XX gastrointestinal-Gen, antisthmatic, antiarteriosclerotic,
XX immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
XX vulnery activity or for gene therapy. The proteins and nucleotides are
XX useful for diagnosing, preventing, treating, or ameliorating conditions
XX or diseases associated with the NF-kappaB pathway. The condition is an
XX immune disorder, an inflammatory disorder, an inflammatory disorder
XX related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
XX hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM
XX syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic
XX ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,
XX hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
XX survival, evasion of immune responses, rheumatoid arthritis, inflammatory
XX bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
XX syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
XX immune activity, disorders related to aberrant acute phase responses,
XX hypercongenital conditions, birth defects, necrotic lesions, wounds,
XX organ transplant rejection, conditions related to organ transplant
XX rejection, disorders related to aberrant signal transduction,
XX proliferating disorders, cancers and HIV propagation in cells infected
XX with other viruses. The present sequence is that of a human protein which
XX is subject to the novel association with the NF-kappaB pathway of the
XX invention. Note: This sequence does not appear in the specification but
XX was obtained by the indexer from Genbank.
XX SQ Sequence 484 AA;
Query Match 78.3%; Score 1887; DB 8; Length 484;
Best Local Similarity 82.1%; Pred. No. 3.3e-171;
Matches 352; Conservative 30; Mismatches 45; Indels 2; Gaps 2;
QY 18 SHYLEARSLNRDVRDYDEYNDYCEGVPRHYHRTBSGYRHCSSKSVSRSSP 77
DB 54 SHYLESRSINEKDYHSRRYIDYERNDYTGCGEPGHRQDRHESRYQNHSSKSGSGSS 113
QY 78 KRK-RNRHCSSH-QSRSKSHRKRKRSSTDEEGHLICQSGDVLRLARYEIVDTLGEAGF 135
DB 114 KSKRIHHTSHRRSHGSHKRRKTRSVDEDEGHLICQSGDVLRLARYEIVDTLGEAGF 173
QY 136 KVECIDHGMDCMHVAVKIVKNGRYEARSEIQVLEHLNSTDPNSVFCVQMLEWFDH 195
DB 174 KVECIDHKGGRHVAVKIVKNDRYCEAARSEIQVLEHLNTDPNSTFRFCVQMLEWFEH 233
QY 196 HGHVCIIVFELLGLSTYDFIKENSFLPFDIHIQWMAQYQICQSNFLHKNKLTHTDLKPEN 255

DB 234 HGHVCIIVFELLGLSTYDFIKENGFLPFDLHIKMAQYQICQSNFLHKNKLTHTDLKPEN 293
QY 256 ILFVKSDYVVKYNSKMKRDERTLKNVDIKVDFGSATYDDEHHSTLVSTRHYRAPEVILA 315
DB 294 ILFQSDYTEAYNPKIKRDERTLINFDIKVDFGSATYDDEHHSTLVSTRHYRAPEVILA 353
QY 316 LGWSQPCDVMSIGCILLIYYLGFVFTQTHDSKEHLAMMERILGPIPOHMIQTKRKRYFH 375
DB 354 LGWSQPCDVMSIGCILLIYYLGFVFTQTHDSKEHLAMMERILGPIPOHMIQTKRKRYFH 413
QY 376 HNOIDWDEHSSAGRYVRRRCCKPLKEPMLCHDEHEKFLDLVRRMLEYDPTQRTILBALQ 435
DB 414 HDBLDWDEHSSAGRYVSRACKPLKEPMLSQDVEHERLFDLIQKQLEYDPAKRITLREALK 473
QY 436 HPFFDLKK 444
DB 474 HPFFDLKK 482

Search completed: March 13, 2005, 00:05:04
Job time : 170 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2005, 14:32:04 ; Search time 1259.64 Seconds
(without alignments)
10662.144 Million cell updates/sec

Title: US-10-801-671-1_COPY_72_2327

Perfect score: 2256

Sequence: 1 gacactcaccagtcattt.....aaacttttgacttcagaaa 2256

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5544816 seqs, 2976611598 residues

Total number of hits satisfying chosen parameters: 11089632

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09D_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2256	100.0	2354	9	US-09-810-671-1
2	2256	100.0	2354	13	US-10-109-854-1
3	2256	100.0	2354	18	US-10-339-656-1
4	1908.4	84.6	4035	17	US-10-801-671-1
5	1286	57.0	1446	17	US-10-425-114-26852
6	1178.2	52.2	1456	17	US-10-267-502-135
7	1097.4	48.6	1549	9	US-10-641-643-699
8	1097.4	48.6	1549	18	US-09-905-999-26
9	1097.2	48.6	1446	17	US-10-825-177-26
10	1097.2	48.6	1446	17	US-10-267-502-138
11	1097	48.6	21234	9	US-09-810-671-3

12	1097	48.6	21234	13	US-10-109-854-3	Sequence 3, Appli
13	1097	48.6	21234	15	US-10-339-656-3	Sequence 3, Appli
14	1097	48.6	21234	18	US-10-801-671-3	Sequence 3, Appli
15	977.4	43.0	3040	17	US-10-425-114-26212	Sequence 26212, A
16	813.2	36.0	1455	17	US-10-267-502-136	Sequence 136, App
17	811.6	36.0	1834	17	US-10-154-708-3	Sequence 3, Appli
18	811.6	36.0	1834	18	US-10-755-889-1	Sequence 1, Appli
19	768.8	34.1	1452	17	US-10-267-502-139	Sequence 139, App
20	768.4	34.1	906	17	US-10-267-502-140	Sequence 140, App
21	644.4	28.6	1743	17	US-10-154-708-10	Sequence 10, Appl
22	640.6	28.4	2254	10	US-09-919-039-238	Sequence 238, App
23	640	28.4	2516	9	US-09-925-298-121	Sequence 121, App
24	640	28.4	2516	14	US-10-102-806-121	Sequence 121, App
25	499.6	22.1	1500	10	US-09-790-852-2	Sequence 2, Appli
26	499.6	22.1	1973	9	US-09-962-436-366	Sequence 266, App
27	499.6	22.1	1973	9	US-09-880-107-2190	Sequence 2190, Ap
28	499.6	22.1	1973	18	US-10-737-450-131	Sequence 131, App
29	481.2	21.3	1538	9	US-09-905-999-22	Sequence 22, Appl
30	481.2	21.3	1538	18	US-10-825-177-22	Sequence 22, Appl
31	464.4	20.6	1296	17	US-10-182-243-16	Sequence 16, Appl
32	464.4	20.6	1628	17	US-10-425-114-26266	Sequence 26266, A
33	461.6	20.5	1026	17	US-10-267-502-133	Sequence 133, App
34	458.6	20.3	1885	17	US-10-439-703-89	Sequence 89, Appl
35	454.8	20.2	1787	9	US-09-905-999-24	Sequence 24, Appl
36	454.8	20.2	1787	18	US-10-825-177-24	Sequence 24, Appl
37	453.2	20.1	1473	17	US-10-267-502-137	Sequence 137, App
38	432.4	19.2	1473	17	US-10-267-502-134	Sequence 134, App
39	432.4	19.2	1762	17	US-10-305-720-1439	Sequence 1439, Ap
40	430.8	19.1	2505	17	US-10-108-260A-2256	Sequence 2256, Ap
41	429.2	19.0	2621	17	US-10-104-047-656	Sequence 656, App
42	377.4	16.7	1181	17	US-10-425-114-16170	Sequence 16170, A
43	369	16.4	369	9	US-09-796-692-5456	Sequence 5456, Ap
44	369	16.4	369	14	US-10-040-862-5456	Sequence 5456, Ap
45	369	16.4	369	17	US-10-057-475B-5456	Sequence 5456, Ap

ALIGNMENTS

RESULT 1

US-09-810-671-1
; Sequence 1, Application US/09810671
; Publication No. US20020076783A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758
; CURRENT APPLICATION NUMBER: US/09/810,671
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Human
US-09-810-671-1

Query Match 100.0%; Score 2256; DB 9; Length 2354;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GACACTCATCCAGTCATTATTAGAACCAAGGTCCTTTGAATGACGAGATTCGGNAC	60
Db	72	GACACTCATCCAGTCATTATTAGAACCAAGGTCCTTTGAATGACGAGATTCGGGAC	131
Qy	61	CGGAGATACGTTTGCGAATACAGGAATCACTACTGTGAAGGATATGTTCCCTAGACATTAT	120
Db	132	CGGAGATACGTTTGCGAATACAGGAATCACTACTGTGAAGGATATGTTCCCTAGACATTAT	191
Qy	121	CACAGAGACATTGAAAGCGGGTATCGAATCCACTGCAGTAAATCTTCAGTCCGACGAGG	180

Db	192	CACAGAGACATTTGAAAGCGGGTATCGAATCCACTTGCAGTAATACTTTCAGTCCGACGACGG	251
Qy	181	AGAAGCAGTCTCTAAAGGAAGCGCAATAGACACACTGTTTCAAGTCATCAGTACAGTTCGAAG	240
Db	252	AGAAGCAGTCTCTAAAGGAAGCGCAATAGACACACTGTTTCAAGTCATCAGTACAGTTCGAAG	311
Qy	241	AGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATGAGGAGGGTCACTGATCTGT	300
Db	312	AGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATGAGGAGGGTCACTGATCTGT	371
Qy	301	CAAAAGTCAGACGTTCTTAAGAGCAAGATATCAAAATCGTGGACATCTTGGGTGAAGGACCC	360
Db	372	CAAAAGTCAGACGTTCTTAAGAGCAAGATATCAAAATCGTGGACATCTTGGGTGAAGGACCC	431
Qy	361	TTTGGCAAAAGTTGTAGAGTGCATTTGATCATGGCATGGATGGCATGTCATGTAGCAGTCAAA	420
Db	432	TTTGGCAAAAGTTGTAGAGTGCATTTGATCATGGCATGGATGGCATGTCATGTAGCAGTCAAA	491
Qy	421	ATCGTAAAAAATGTATAGGCGGTTTACCGTGAAAGAGCTCGTTCAGAAAATCCAAAGTATTAGAG	480
Db	492	ATCGTAAAAAATGTATAGGCGGTTTACCGTGAAAGAGCTCGTTCAGAAAATCCAAAGTATTAGAG	551
Qy	481	CACTTTAAATAGTACTGATCCCAATAGTGTCTTCCGATGTGCCAGATGCTCAGATGCTTT	540
Db	552	CACTTTAAATAGTACTGATCCCAATAGTGTCTTCCGATGTGCCAGATGCTCAGATGCTTT	611
Qy	541	GATCATCATGTGTCATGTTGTATTGTTGAACTACTTGGGACCTTAGTACTTACGATTTTC	600
Db	612	GATCATCATGTGTCATGTTGTATTGTTGTTGAACTACTTGGGACCTTAGTACTTACGATTTTC	671
Qy	601	ATTAAAGAAAAACAGCTTCTGCCATTTCCAAATGACCAATCAGGCAGATGGCGTATCAG	660
Db	672	ATTAAAGAAAAACAGCTTCTGCCATTTCCAAATGACCAATCAGGCAGATGGCGTATCAG	731
Qy	661	ATCTGCCAGTCAATAAATTTTTTACATCATATAAATTTAAATTAACCCATACAGATCTGAAGCCT	720
Db	732	ATCTGCCAGTCAATAAATTTTTTACATCATATAAATTTAAATTAACCCATACAGATCTGAAGCCT	791
Qy	721	GAAAAATATTTTGTGTTGAAAGTCTGACATGTGCTCAATATAATTTCTTAAATCAAAACGT	780
Db	792	GAAAAATATTTTGTGTTGAAAGTCTGACATGTGCTCAATATAATTTCTTAAATCAAAACGT	851
Qy	781	GATGAACGCCACCTGAAAAAACACAGATATCAAAAGTTGTGTGCTTTGGAAGTGCACACGTAT	840
Db	852	GATGAACGCCACCTGAAAAAACACAGATATCAAAAGTTGTGTGCTTTGGAAGTGCACACGTAT	911
Qy	841	GATGATGAACATCACAGTACTTTGGTGTCTACCCGGCACTACAGAGTCCCGAGGTCAATT	900
Db	912	GATGATGAACATCACAGTACTTTGGTGTCTACCCGGCACTACAGAGTCCCGAGGTCAATT	971
Qy	901	TTGGCTTTTAGTTGGTCTCAGCCTTGATGTTTGGAGCATAGGTTGCACTTTATTGAA	960
Db	972	TTGGCTTTTAGTTGGTCTCAGCCTTGATGTTTGGAGCATAGGTTGCACTTTATTGAA	1031
Qy	961	TATTACCTTTGGTTTCAAGTCTTTCAGATCATGATAGTAAGAGCACCTGGGCAATGATG	1020
Db	1032	TATTACCTTTGGTTTCAAGTCTTTCAGATCATGATAGTAAGAGCACCTGGGCAATGATG	1091
Qy	1021	GAAACGAATATTAGACCCATACCAACACATGATTTAGAAAAACAGAAAAACGCAAGTAT	1080
Db	1092	GAAACGAATATTAGACCCATACCAACACATGATTTAGAAAAACAGAAAAACGCAAGTAT	1151
Qy	1081	TTTTCACCAATAACAGTATGTTGGATGAAACACACAGTTTCTGCTGGTATGATATGTTAGAGA	1140
Db	1152	TTTTCACCAATAACAGTATGTTGGATGAAACACACAGTTTCTGCTGGTATGATATGTTAGAGA	1211
Qy	1141	CGCTGCAAAACCGTTGAAGGAATTTTATGCTTTTGTATGATAGTAAGAACATGAGAAAACGTGTTT	1200
Db	1212	CGCTGCAAAACCGTTGAAGGAATTTTATGCTTTTGTATGATAGTAAGAACATGAGAAAACGTGTTT	1271
Qy	1201	GACCTGGTTTCGAAGAAATGTTTATGATCAACTCAAGAAATTAACCTTGGATGAAGACA	1260
Db	1272	GACCTGGTTTCGAAGAAATGTTTATGATCAACTCAAGAAATTAACCTTGGATGAAGACA	1331

Qy	1261	TTGCAGCATCCTTTCTTTGTGACTTATTAATAAGAAATGAATCGGATCAGTGGTCTTAC	1320
Db	1332	TTGCAGCATCCTTTCTTTGTGACTTATTAATAAGAAATGAATCGGATCAGTGGTCTTAC	1391
Qy	1321	TATATACCTCTCTAGAAGAGATTAACCTTAAGAATGTGTGTCAGTCAACTAAACATTTCTAATAT	1380
Db	1392	TATATACCTCTCTAGAAGAGATTAACCTTAAGAATGTGTGTCAGTCAACTAAACATTTCTAATAT	1451
Qy	1381	TTTTGTAAACATTAATAATTTTTGTACAGTTAAGTGTAAATATGTATGTTTGTATCAA	1440
Db	1452	TTTTGTAAACATTAATAATTTTTGTACAGTTAAGTGTAAATATGTATGTTTGTATCAA	1511
Qy	1441	TAGCATAAATTAACCTTGTTTAAGCAAGTATGGTCTTTGATAAATGCATTAAGAAAAATTAATAAT	1500
Db	1512	TAGCATAAATTAACCTTGTTTAAGCAAGTATGGTCTTTGATAAATGCATTAAGAAAAATTAATAAT	1571
Qy	1501	AATTTTTCTTTTGAANAATACCATTTTTTAATACCTTTGAANAATCTCTTTGTGTCCAGTG	1560
Db	1572	AATTTTTCTTTTGAANAATACCATTTTTTAATACCTTTGAANAATCTCTTTGTGTCCAGTG	1631
Qy	1561	ATAAATGTGATGTGATCTTTGCTTTGTACATGGAGGTCACTCTGAAGTGATTTTTTTTTTG	1620
Db	1632	ATAAATGTGATGTGATCTTTGCTTTGTACATGGAGGTCACTCTGAAGTGATTTTTTTTTTG	1691
Qy	1621	AGTAAAGGAAATCTTGACCTACTTTATATTCTTTAAAGGAATATCTTTATATACTTCAAA	1680
Db	1692	AGTAAAGGAAATCTTGACCTACTTTATATTCTTTAAAGGAATATCTTTATATACTTCAAA	1751
Qy	1681	TTTAGAACCTTAACTTTTAAAAAGTTTTCTTCTGTAAATTTGTTGAAACGGGTGATTAATTAA	1740
Db	1752	TTTAGAACCTTAACTTTTAAAAAGTTTTCTTCTGTAAATTTGTTGAAACGGGTGATTAATTAA	1811
Qy	1741	CTCTAGATAGCAGGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTGTAGAAATCTTAT	1800
Db	1812	CTCTAGATAGCAGGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTGTAGAAATCTTAT	1871
Qy	1801	TAAATTTTAAAGTGTGTATTTCTTTTTCATTTGGGTGATGTCAGGGTGATAACACAGACATTC	1860
Db	1872	TAAATTTTAAAGTGTGTATTTCTTTTTCATTTGGGTGATGTCAGGGTGATAACACAGACATTC	1931
Qy	1861	ATGAAAGGCAATGACAGTTTGTCCATTTGTGACAGTTTGTAAATAAACCAATACACACT	1920
Db	1932	ATGAAAGGCAATGACAGTTTGTCCATTTGTGACAGTTTGTAAATAAACCAATACACACT	1991
Qy	1921	TTATTTTAAGATTAAATCTAACTGGAAGTCAGCTTGGAAAAATGGACATTTCCAAAGTATG	1980
Db	1992	TTATTTTAAGATTAAATCTAACTGGAAGTCAGCTTGGAAAAATGGACATTTCCAAAGTATG	2051
Qy	1981	TTTGGTGAGTCACAGATATAAAAATAGAAATTTCTGATGAGAGGTTTCAGTTTTTAAATACC	2040
Db	2052	TTTGGTGAGTCACAGATATAAAAATAGAAATTTCTGATGAGAGGTTTCAGTTTTTAAATACC	2111
Qy	2041	AAGTCTTTTAGAGTCCTAAACATTTGGCCAGCATCTGTTTTATCAAAATGACATAAATACGTAA	2100
Db	2112	AAGTCTTTTAGAGTCCTAAACATTTGGCCAGCATCTGTTTTATCAAAATGACATAAATACGTAA	2171
Qy	2101	ACCTATAAGAAATTAAGTTTATTAATATAGGCAATTTATGTCTGTGATAAATTTCTTACGGGAG	2160
Db	2172	ACCTATAAGAAATTAAGTTTATTAATATAGGCAATTTATGTCTGTGATAAATTTCTTACGGGAG	2231
Qy	2161	AAAGAGGATTTGATTTGGAACGAGTTTGGGAAGAAAGTGTGCTGAGAAATTTTCCAGAAATTT	2220
Db	2232	AAAGAGGATTTGATTTGGAACGAGTTTGGGAAGAAAGTGTGCTGAGAAATTTTCCAGAAATTT	2291
Qy	2221	AAATTGATGGTTACATAAATCTTTTGTACTTTCAGAAA	2256
Db	2292	AAATTGATGGTTACATAAATCTTTTGTACTTTCAGAAA	2327

```
; Publication No. US20020119548A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758DIV
; CURRENT APPLICATION NUMBER: US/10/109,854
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-109-854-1

Query Match      100.0%; Score 2256; DB 13; Length 2354;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACTCATCCAGTCATTATTAGAGCAAGGTCCTTGAATGAGCGAGATTATCGGAC 60
DB 72 GACACTCATCCAGTCATTATTAGAGCAAGGTCCTTGAATGAGCGAGATTATCGGAC 131

QY 61 CGGAGATACGTTGAGGATACAGGATGACTACTGTGAGGATATGTCCTAGACATTAT 120
DB 132 CGGAGATACGTTGAGGATACAGGATGACTACTGTGAGGATATGTCCTAGACATTAT 191

QY 121 CACAGAGACATTTGAAGCGGGTATCGAATCCACTGCAGTAAATCTTCAGTCGCGAGCAGG 180
DB 192 CACAGAGACATTTGAAGCGGGTATCGAATCCACTGCAGTAAATCTTCAGTCGCGAGCAGG 251

QY 181 AGAAGCAGTCCTTAAAGGAAGCGCAATAGACACACTGTTCAAGTCATCAGTCACGTTTCGAAG 240
DB 252 AGAAGCAGTCCTTAAAGGAAGCGCAATAGACACACTGTTCAAGTCATCAGTCACGTTTCGAAG 311

QY 241 AGCCACCAAGGAAAGATCCAGGATGATAGGATGATAGGAGGGTCACCTGATCTGT 300
DB 312 AGCCACCAAGGAAAGATCCAGGATGATAGGATGATAGGAGGGTCACCTGATCTGT 371

QY 301 CAAGTGGAGAGCTTCTAAGAGCAAGATATGAATCGTGGACACTTTGGGTGAAGAGCC 360
DB 372 CAAGTGGAGAGCTTCTAAGAGCAAGATATGAATCGTGGACACTTTGGGTGAAGAGCC 431

QY 361 TTGSCAAAGTTGTAGAGTGATTCATGATCGCATGGATGGCATGCAATGTAGCAGTGAAA 420
DB 432 TTGSCAAAGTTGTAGAGTGATTCATGATCGCATGGATGGCATGCAATGTAGCAGTGAAA 491

QY 421 ATCGTAAAGATGATAGGCGGTTACCGTGAAGCAGCTCGTTGAGAAATCCAAGTATTAGAG 480
DB 492 ATCGTAAAGATGATAGGCGGTTACCGTGAAGCAGCTCGTTGAGAAATCCAAGTATTAGAG 551

QY 481 CACTTAAATAGTACTGATCCCAATAGTGTCTCCGATGTGTCAGATGCTAGAGTGGTTT 540
DB 552 CACTTAAATAGTACTGATCCCAATAGTGTCTCCGATGTGTCAGATGCTAGAGTGGTTT 611

QY 541 GATCATCATGTCATGTTTGTATGTTGTTGAATCTAGTGGGACTTGTAGTACTTACGATTTTC 600
DB 612 GATCATCATGTCATGTTTGTATGTTGTTGAATCTAGTGGGACTTGTAGTACTTACGATTTTC 671

QY 601 ATTAAGAAAACAGCTTTCTGCGCATTTCAAATGACCAATCAGGCAGATGCGGTATCAG 660
DB 672 ATTAAGAAAACAGCTTTCTGCGCATTTCAAATGACCAATCAGGCAGATGCGGTATCAG 731

QY 661 ATCTGCCAGTCAATAAATTTTTTACATCATATAAATTAACCCATACAGATCTGAAGCCT 720
DB 732 ATCTGCCAGTCAATAAATTTTTTACATCATATAAATTAACCCATACAGATCTGAAGCCT 791
```

Db 1872 TAAATTTTAAAGTGTGTATCTTTTTCATGCGGTGATGTCAGGGTGTAAACCGAGCATTC 1931
Qy 1861 ATGGAAGGCGATGCAAGTTTGTCCATTTGACAGTTTGTATTAATAAACCACATACACACT 1920
Db 1932 ATGGAAGGCGATGCAAGTTTGTCCATTTGACAGTTTGTATTAATAAACCACATACACACT 1991
Qy 1921 TTATTTAAGATTTAAATCTAACTGGAAGTCAAGCTTGGAAATGACATTTCCAAAGTATG 1980
Db 1932 TTATTTAAGATTTAAATCTAACTGGAAGTCAAGCTTGGAAATGACATTTCCAAAGTATG 2051
Qy 1981 TTGTGTAGTACAGATATAAATAAATAGAAATTTCTGATGAGAGGTTTCAGTTTAAATACC 2040
Db 2052 TTGTGTAGTACAGATATAAATAAATAGAAATTTCTGATGAGAGGTTTCAGTTTAAATACC 2111
Qy 2041 AAGTCCTTAGGAGTCTTAACATTTGGCCAGCAGCTCTGTTTCAATCAATGACATAAATACGTAA 2100
Db 2112 AAGTCCTTAGGAGTCTTAACATTTGGCCAGCAGCTCTGTTTCAATCAATGACATAAATACGTAA 2171
Qy 2101 ACCTATAAGAAATTAAGTTTATTAATTAAGGCAATTTATGTCGTGATAATTTCTTACGGGAG 2160
Db 2172 ACCATATAAGATTAAGTTTATTAATTAAGGCAATTTATGTCGTGATAATTTCTTACGGGAG 2231
Qy 2161 AAAGAGATTTGATTTGGAAGCAGTTTGGGAAGAAAGTGTCTGCTGAAATTTCCAGAAATTT 2220
Db 2232 AAAGAGATTTGATTTGGAAGCAGTTTGGGAAGAAAGTGTCTGCTGAAATTTCCAGAAATTT 2291
Qy 2221 AATTGATTTGGTTACATAAATCTTTTGAATCTCAGAAA 2256
Db 2292 AATTGATTTGGTTACATAAATCTTTTGAATCTCAGAAA 2327

RESULT 3

US-10-339-656-1
; Sequence 1, Application US/10339656
; Publication No. US20030134319A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758DIV2
; CURRENT APPLICATION NUMBER: US/10/339,656
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/109,854
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-339-656-1

Query Match 100.0%; Score 2256; DB 15; Length 2354;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACACTCATCCAGTCAATTTATTAAGCAAGGTCCTTTGAATGAGCGAGATTATCGGAC 60
Db 72 GACACTCATCCAGTCAATTTATTAAGCAAGGTCCTTTGAATGAGCGAGATTATCGGAC 131
Qy 61 CGGAGATCGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTAT 120
Db 132 CGGAGATCGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTAT 191
Qy 121 CACAGAGATTTGAAGCGGGTATCGAATCCACTCCAGTAAATCTTCAGTCGGAGCAGG 180
Db 192 CACAGAGATTTGAAGCGGGTATCGAATCCACTCCAGTAAATCTTCAGTCGGAGCAGG 251

Qy 181 AGAAGCAGTCTCTAAAGAGGAGCGCAATAGACACTGTTTCAAGTCATCAGTCACTGTTCTGAAG 240
Db 252 AGAAGCAGTCTCTAAAGAGGAGCGCAATAGACACTGTTTCAAGTCATCAGTCACTGTTCTGAAG 311
Qy 241 AGCCACCGAAGGAAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGT 300
Db 312 AGCCACCGAAGGAAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGT 371
Qy 301 CAAGTGTGAGACGTTCTTAAGAGCAAGATATGAAATCGTGGACACTTTGGGTGAAGAGGCC 360
Db 372 CAAGTGTGAGACGTTCTTAAGAGCAAGATATGAAATCGTGGACACTTTGGGTGAAGAGGCC 431
Qy 361 TTTGGCAAAAGTTGTAGAGTGCATTTGATCATGCGCATGGATGCGCATGTAGCAGTGAA 420
Db 432 TTTGGCAAAAGTTGTAGAGTGCATTTGATCATGCGCATGGATGCGCATGTAGCAGTGAA 491
Qy 421 ATCGTAAAAATGTAGGCGGTTACCGTGAAGCAGCTGTTCCAGAAAATCCAAAGTATTAGAG 480
Db 492 ATCGTAAAAATGTAGGCGGTTACCGTGAAGCAGCTGTTCCAGAAAATCCAAAGTATTAGAG 551
Qy 481 CACTTAAATAGTACTGATCCCAATAGTGTCTTCCGATGTGCCAGATGCTAGAAATGGTTT 540
Db 552 CACTTAAATAGTACTGATCCCAATAGTGTCTTCCGATGTGCCAGATGCTAGAAATGGTTT 611
Qy 541 GATCATCATGTCATGTTTGTATTTGTTGAACTACTGGGACTTAGTACTTACGATTTTC 600
Db 612 GATCATCATGTCATGTTTGTATTTGTTGAACTACTGGGACTTAGTACTTACGATTTTC 671
Qy 601 ATTAAGAAAACAGCTTCTGCCAATTCAAATGACCAATCAGGAGATGCGGTATCAG 660
Db 672 ATTAAGAAAACAGCTTCTGCCAATTCAAATGACCAATCAGGAGATGCGGTATCAG 731
Qy 661 ATCTGCCAGTCAATAAATTTTTTACATCATATAAATTAACCCATACAGATCTGAAGCCT 720
Db 732 ATCTGCCAGTCAATAAATTTTTTACATCATATAAATTAACCCATACAGATCTGAAGCCT 791
Qy 721 GAAAATATTTTGTGTGAAGTCTGACTATGCTGAAATATAAATTTTAAATGAAACGT 780
Db 792 GAAAATATTTTGTGTGAAGTCTGACTATGCTGAAATATAAATTTTAAATGAAACGT 851
Qy 781 GATGAACGACACTGAAAACACAGATATCAAGTTGTGTGACTTTGGAGTGCACAGCTAT 840
Db 852 GATGAACGACACTGAAAACACAGATATCAAGTTGTGTGACTTTGGAGTGCACAGCTAT 911
Qy 841 GATGATGAACATCAAGTACTTTGCTGTCTACCGGCACTACAGAGCTCCCGAGGTCATT 900
Db 912 GATGATGAACATCAAGTACTTTGCTGTCTACCGGCACTACAGAGCTCCCGAGGTCATT 971
Qy 901 TTGGCTTTTAGGTTGGTCTCAGCCTTTGTGATGTTTGGAGCATAGGTTGCATTTATTGAA 960
Db 972 TTGGCTTTTAGGTTGGTCTCAGCCTTTGTGATGTTTGGAGCATAGGTTGCATTTATTGAA 1031
Qy 961 TATTACCTTTGTTTCAAGTCTTTTCAAGTCTCATGATAGTAAAGACACCTGGCAATGATG 1020
Db 1032 TATTACCTTTGTTTCAAGTCTTTTCAAGTCTCATGATAGTAAAGACACCTGGCAATGATG 1091
Qy 1021 GAACGAATATTAGGACCCATACCAACACATGATTTTCAAGAAAACAAAGAAACGCAAGTAT 1080
Db 1092 GAACGAATATTAGGACCCATACCAACACATGATTTTCAAGAAAACAAAGAAACGCAAGTAT 1151
Qy 1081 TTTTCAACATAACCAGTATAGTTGGGATGAACACAGTTTCTGCTGTGTAGATATGTTAGGAGA 1140
Db 1152 TTTTCAACATAACCAGTATAGTTGGGATGAACACAGTTTCTGCTGTGTAGATATGTTAGGAGA 1211
Qy 1141 CGCTGCAAAACGTTGAAGGAATTTATGCTTTTGTATGATGAAGACATGAGAAATGTTT 1200
Db 1212 CGCTGCAAAACGTTGAAGGAATTTATGCTTTTGTATGATGAAGAAACATGAGAAATGTTT 1271
Qy 1201 GACCTGGTTTGAAGAAATGTTTAGAATATGATCAACTCAAGAAATTTACCTTGGATGAAGCA 1260
Db 1272 GACCTGGTTTGAAGAAATGTTTAGAATATGATCAACTCAAGAAATTTACCTTGGATGAAGCA 1331

QY 1261 TTGAGGATCTCTTCTTGAATTATAAAGAAATGAATGGGAATCAGTGGCTTAC 1320
 Db 1332 TTGAGGATCTCTTCTTGAATTATAAAGAAATGAATGGGAATCAGTGGCTTAC 1391
 QY 1321 TATATACCTCTCTAGAGAGATTACTTAAAGCTGTGCTAGTCAATCAATCTCAATAT 1380
 Db 1392 TATATACCTCTCTAGAGAGATTACTTAAAGCTGTGCTAGTCAATCAATCTCAATAT 1451
 QY 1381 TTTTGTAAACATTAATATTTTGTACAGTGTAAATATTTATTTTGTATCAAA 1440
 Db 1452 TTTTGTAAACATTAATATTTTGTACAGTGTAAATATTTATTTTGTATCAAA 1511
 QY 1441 TAGCAATATTAATCTTTTAAAGAGATGCTCTGATTAATGCAATAGAAAATTTAAAT 1500
 Db 1512 TAGCAATATTAATCTTTTAAAGAGATGCTCTGATTAATGCAATAGAAAATTTAAAT 1571
 QY 1501 AATTTTCTTTTGAATTTACCATTTTAAATACCTTTGAAATATCCTTTGTGCCAGT 1560
 Db 1572 AATTTTCTTTTGAATTTACCATTTTAAATACCTTTGAAATATCCTTTGTGCCAGT 1631
 QY 1561 ATAAATGTGATTCCTTTCCTTTTGTATCATGGAGTCACTCTCAAGTGAATTTTGTG 1620
 Db 1632 ATAAATGTGATTCCTTTCCTTTTGTATCATGGAGTCACTCTCAAGTGAATTTTGTG 1691
 QY 1621 AGTAAAGGAAATCTTGACTACTTTATATTTCTTAAAGGAATATCTTTATATACTTCAA 1680
 Db 1692 AGTAAAGGAAATCTTGACTACTTTATATTTCTTAAAGGAATATCTTTATATACTTCAA 1751
 QY 1681 TTTAGAACTTAACCTTTAAAGTTTCTTCTGTAATCTTGAACGGGTGATTTATTTAA 1740
 Db 1752 TTTAGAACTTAACCTTTAAAGTTTCTTCTGTAATCTTGAACGGGTGATTTATTTAA 1811
 QY 1741 CTCTAGATAAGCAGTACTAGAAAACCAAACTCAGAAAATGTTTACTGTAGAAATTCAT 1800
 Db 1812 CTCTAGATAAGCAGTACTAGAAAACCAAACTCAGAAAATGTTTACTGTAGAAATTCAT 1871
 QY 1801 TAAATTTTAAAGTTTGTATCTTTTTCATTTGGTGTAGTCAAGGTGATAACAGACATTC 1860
 Db 1872 TAAATTTTAAAGTTTGTATCTTTTTCATTTGGTGTAGTCAAGGTGATAACAGACATTC 1931
 QY 1861 ATGGAAGGCAATGCAAGTTTGTCTTCTGATTTGACAGTTTGTATTAATAAACCACATACACT 1920
 Db 1932 ATGGAAGGCAATGCAAGTTTGTCTTCTGATTTGACAGTTTGTATTAATAAACCACATACACT 1991
 QY 1921 TTAATTAAGATTAATCTAACTGAAAGTCAAGTGTGAAATGGAATTTTCAAGTATG 1980
 Db 1992 TTAATTAAGATTAATCTAACTGAAAGTCAAGTGTGAAATGGAATTTTCAAGTATG 2051
 QY 1981 TTTGGTGTAGTCAAGATATAAATAAGAAATTTCTGATGAGAGTTTCAAGTTTAAATACC 2040
 Db 2052 TTTGGTGTAGTCAAGATATAAATAAGAAATTTCTGATGAGAGTTTCAAGTTTAAATACC 2111
 QY 2041 AAGTCTTTAGAGTTTAACTTGGCCAGCATCTGTTTATCAATGACATAAATACGTAA 2100
 Db 2112 AAGTCTTTAGAGTTTAACTTGGCCAGCATCTGTTTATCAATGACATAAATACGTAA 2171
 QY 2101 ACCTATAAGATTAAGTTTATTAATTTAGCAATTTATCTGTGATTAATTTCTAGCGGAG 2160
 Db 2172 ACCTATAAGATTAAGTTTATTAATTTAGCAATTTATCTGTGATTAATTTCTAGCGGAG 2231
 QY 2161 AAAGAGATTTGATTTGAAAGCAGTTTGGGAAGAAAGTGTCTGCTGAAATTTCCAGAAATTT 2220
 Db 2232 AAAGAGATTTGATTTGAAAGCAGTTTGGGAAGAAAGTGTCTGCTGAAATTTCCAGAAATTT 2291
 QY 2221 AATTTGATTTGATTAATAACTTTTGTACTTCAGAAA 2256
 Db 2292 AATTTGATTTGATTAATAACTTTTGTACTTCAGAAA 2327

RESULT 4

US-10-801-671-1
 ; Sequence 1, Application US/10801671
 ; Publication No. US20040152123A1

; GENERAL INFORMATION:
 ; APPLICANT: YAN, Chunhua et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: CL000758DIV-III
 ; CURRENT APPLICATION NUMBER: US/10/801,671
 ; CURRENT FILING DATE: 2004-03-17
 ; PRIOR APPLICATION NUMBER: 60/227,470
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: 09/810,671
 ; PRIOR FILING DATE: 2001-03-19
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 2354
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-801-671-1

Query Match 100.0%; Score 2256; DB 18; Length 2354;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACTCATCCAGTCATTATTTAGAACAAAGGTCCTTGAATGAGCGAGATTTATCGGAC 60
 Db 72 GACACTCATCCAGTCATTATTTAGAACAAAGGTCCTTGAATGAGCGAGATTTATCGGAC 131
 QY 61 CGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGATATGTTCTTAGACATTTAT 120
 Db 132 CGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGATATGTTCTTAGACATTTAT 191
 QY 121 CACAGAGACATTTGAAGCGGTATTCGAATCCACTGCAGTAAATCTTCACTCGCAGCAGG 180
 Db 192 CACAGAGACATTTGAAGCGGTATTCGAATCCACTGCAGTAAATCTTCACTCGCAGCAGG 251
 QY 181 AGAAGCAGTCTTAAAGAAAGCGCAATAGACACTGTTCAAGTCACTCAGTCAGTTCGAAG 240
 Db 252 AGAAGCAGTCTTAAAGAAAGCGCAATAGACACTGTTCAAGTCACTCAGTCAGTTCGAAG 311
 QY 241 AGCCACCGAAGAAAGATCCAGGATATAGAGATGATGAGGAGGTCACCTGATCTGT 300
 Db 312 AGCCACCGAAGAAAGATCCAGGATATAGAGATGATGAGGAGGTCACCTGATCTGT 371
 QY 301 CAAAGTGGAGACGTTCTTAAGAGCAAGATATGAATCGTGGACACTTTGGGTGAAGAGCC 360
 Db 372 CAAAGTGGAGACGTTCTTAAGAGCAAGATATGAATCGTGGACACTTTGGGTGAAGAGCC 431
 QY 361 TTTGGCAAAAGTTGAGAGTGCATTTGATCATGGCATGGATGGCATGTCATGTGACGTGAAA 420
 Db 432 TTTGGCAAAAGTTGAGAGTGCATTTGATCATGGCATGGATGGCATGTCATGTGACGTGAAA 491
 QY 421 ATCGTAAATAATGATAGCGGTTACCGTGAACAGCTCGTTCAGAAATCCAAGTATTAGAG 480
 Db 492 ATCGTAAATAATGATAGCGGTTACCGTGAACAGCTCGTTCAGAAATCCAAGTATTAGAG 551
 QY 481 CACTTAAATAGTACTGATCCCAATAGTCTTCCGATGTGTCGATGTCCGATGTCCGATGTGTT 540
 Db 552 CACTTAAATAGTACTGATCCCAATAGTCTTCCGATGTGTCGATGTCCGATGTGTT 611
 QY 541 GATCATCATGTCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 600
 Db 612 GATCATCATGTCATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 671
 QY 601 ATTAAGAAACACGTTTCTGCCATTTCAATTTGACCAATCAGGCATCAGGCATGCGGTATCAG 660
 Db 672 ATTAAGAAACACGTTTCTGCCATTTCAATTTGACCAATCAGGCATCAGGCATGCGGTATCAG 731
 QY 661 ATCTGCCAGTCAATAAATTTTTTACATCATATAAATTTAAATTAATTAATTAATTAATTAATTAAT 720
 Db 732 ATCTGCCAGTCAATAAATTTTTTACATCATATAAATTTAAATTAATTAATTAATTAATTAATTAAT 791
 QY 721 GAAAAATATTTTGTGTGAGTCTGACTATGCTAGTCAATATATAATTTCTAAAAATGAAACGT 780

Db 792 GAAAAATATTTTGTGTGGAAGCTGACATGATGTAGTCAAAATATAATCTTAAAAATGAACGTT 851
Qy 781 GATGAACGACACTGGAAGAAACACAGATATCAAGTGTGTTGATCTTGGAGTGCACGTAT 840
Db 852 GATGAACGACACTGGAAGAAACACAGATATCAAGTGTGTTGATCTTGGAGTGCACGTAT 911
Qy 841 GATGATGAACATCAAGTACTTGTGTGTTCCCGGCACTACAGAGTCCCGAGGTCAAT 900
Db 912 GATGATGAACATCAAGTACTTGTGTGTTCCCGGCACTACAGAGTCCCGAGGTCAAT 971
Qy 901 TTGGCTTTAGGTGTGTTCTCAGCTTGTGATGTTTGGAGCATAGGTGTGATCTTATTGAA 960
Db 972 TTGGCTTTAGGTGTGTTCTCAGCTTGTGATGTTTGGAGCATAGGTGTGATCTTATTGAA 1031
Qy 961 TATTACCTTGGTTTCAAGTCTTTCAGACTCATGATAGTAAAGAGCACCCTGGCAATGATG 1020
Db 1032 TATTACCTTGGTTTCAAGTCTTTCAGACTCATGATAGTAAAGAGCACCCTGGCAATGATG 1091
Qy 1021 GAAACGAATATTAGGACCCATACCAACACACATGATTCAAGAAACAAAGAAACGCAAGTAT 1080
Db 1092 GAACGAATATTAGGACCCATACCAACACACATGATTCAAGAAACAAAGAAACGCAAGTAT 1151
Qy 1081 TTTCAACCAATACCAAGTATAGTTGGATGAACACAGTCTCTGCTGTGATGATGTTAGGAGA 1140
Db 1152 TTTCAACCAATACCAAGTATAGTTGGATGAACACAGTCTCTGCTGTGATGATGTTAGGAGA 1211
Qy 1141 CGCTGCAAAACCGTTGAAGGAATTTATGCTTTGTCAATGATGAAGAACATGAGAACTGTTT 1200
Db 1212 CGCTGCAAAACCGTTGAAGGAATTTATGCTTTGTCAATGATGAAGAACATGAGAACTGTTT 1271
Qy 1201 GACCTGGTTTGAAGGAATGTTAGAAATATGATCAACTCAAGAAATTAACCTTGGATGAAGCA 1260
Db 1272 GACCTGGTTTGAAGGAATGTTAGAAATATGATCAACTCAAGAAATTAACCTTGGATGAAGCA 1331
Qy 1261 TTGCAGCATCTCTTCTTTGACTTATTAAGAAAGAAATGAAATGGAATCAGTGGTCTTAC 1320
Db 1332 TTGCAGCATCTCTTCTTTGACTTATTAAGAAAGAAATGAAATGGAATCAGTGGTCTTAC 1391
Qy 1321 TATATACCTCTAGAACAGATATCTTAAGACTGTGTGATGATGATCACTAAACATCTTAATAT 1380
Db 1392 TATATACCTCTAGAACAGATATCTTAAGACTGTGTGATGATGATCACTAAACATCTTAATAT 1451
Qy 1381 TTTTGTAAACATTAATATTTTGTACAGTAAATGATGATGATGATGATGATGATGATGATGAT 1440
Db 1452 TTTTGTAAACATTAATATTTTGTACAGTAAATGATGATGATGATGATGATGATGATGATGAT 1511
Qy 1441 TAGCATAAATTAACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Db 1512 TAGCATAAATTAACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1571
Qy 1501 AATTTTCTTTTGAATTAACCTTTTGAATTAACCTTTTGAATTAACCTTTTGAATTAACCTTT 1560
Db 1572 AATTTTCTTTTGAATTAACCTTTTGAATTAACCTTTTGAATTAACCTTTTGAATTAACCTTT 1631
Qy 1561 ATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
Db 1632 ATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1691
Qy 1621 AGTAAAGGAATCTGATCTTATATCTTAAAGGAATATCTTATATATATATATATATATATAT 1680
Db 1692 AGTAAAGGAATCTGATCTTATATCTTAAAGGAATATCTTATATATATATATATATATATAT 1751
Qy 1681 TTTAGAACTTAACCTTTAAAGGTTTCTTCTGATGATGATGATGATGATGATGATGATGATGAT 1740
Db 1752 TTTAGAACTTAACCTTTAAAGGTTTCTTCTGATGATGATGATGATGATGATGATGATGATGAT 1811
Qy 1741 CTCTAGATAAGCAGTACTAGAAACCAAACTCAGAAATGTTTACTGTAGAAATCTTAT 1800
Db 1812 CTCTAGATAAGCAGTACTAGAAACCAAACTCAGAAATGTTTACTGTAGAAATCTTAT 1871
Qy 1801 TAAATTTTAAAGTGTGATCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1860

Db 1872 TAAATTTTAAAGTGTGATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1931
Qy 1861 ATGAAAGGCGATGAGTGTGTCATGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTG 1920
Db 1932 ATGAAAGGCGATGAGTGTGTCATGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTG 1991
Qy 1921 TTAATTTAAGATTTAAATCTTAACCTGGAAGTCACTGGAAGTCACTGGAAGTCACTGGAAGTCACT 1980
Db 1992 TTAATTTAAGATTTAAATCTTAACCTGGAAGTCACTGGAAGTCACTGGAAGTCACTGGAAGTCACT 2051
Qy 1981 TTTGGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040
Db 2052 TTTGGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2111
Qy 2041 AAGTCTTCTAGGAGTCTTAACATTCGCCAGCATCTGTTTATCAATGACATGACATGACATGACATG 2100
Db 2112 AAGTCTTCTAGGAGTCTTAACATTCGCCAGCATCTGTTTATCAATGACATGACATGACATGACATG 2171
Qy 2101 ACCTATTAAGATTTAAGTCTTAAATTAAGTCTTAAATTAAGTCTTAAATTAAGTCTTAAATTAAG 2160
Db 2172 ACCTATTAAGATTTAAGTCTTAAATTAAGTCTTAAATTAAGTCTTAAATTAAGTCTTAAATTAAG 2231
Qy 2161 AAGAGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220
Db 2232 AAGAGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2291
Qy 2221 AATTTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2256
Db 2292 AATTTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2327

RESULT 5

US-10-425-114-26852
; Sequence 26852, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26852
; LENGTH: 4035
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4654-056-G4_FLI
US-10-425-114-26852

Query Match 84.6%; Score 1908.4; DB 17; Length 4035;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1909; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 327 ATATGAATTCGTGGACACTTTGGGTGAAGAGAGCCCTTTGGCAAAAGTTGTAGAGTGCAATGA 386
Db 2126 AGATGAATTCGTGGACACTTTGGGTGAAGAGAGCCCTTTGGCAAAAGTTGTAGAGTGCAATGA 2185
Qy 387 TCATGGCATGATGGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 446
Db 2186 TCATGGCATGATGGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2245
Qy 447 TGAAGCAGCTCGTTTCAAGAAATCCCAAGTATTAGAGCAGCTTAAATAGTACTGATCCCAATAG 506
Db 2246 TGAAGCAGCTCGTTTCAAGAAATCCCAAGTATTAGAGCAGCTTAAATAGTACTGATCCCAATAG 2305
Qy 507 TGTCTTCGATGTGTGTCAGATGCTAGAAATGCTGATGATGATGATGATGATGATGATGATGATGAT 566

```
Db 2306 TGTCTTCCGATGTGTCAGATGCTAGATGGTTTGATCATCATGTGTCATGTTGTTATGTT 2365
Qy 567 GTTTGAACTACTCGGACCTTAGTACTTACGATTTTCAATAAGAAAACAGCTTCTGCGCAAT 626
Db 2366 GTTTGAACTACTCGGACCTTAGTACTTACGATTTTCAATAAGAAAACAGCTTCTGCGCAAT 2425
Qy 627 TCAATTCACCACTACGAGCAGATGGCGTATCAGATCTGCGAGTCAATAAATTTTTCACA 686
Db 2426 TCAATTCACCACTACGAGCAGATGGCGTATCAGATCTGCGAGTCAATAAATTTTTCACA 2485
Qy 687 TCATAATAAATAAATTAATCTTAAATGAAACGTCGATGAACGCACTGAAATAACACAGA 746
Db 2486 TCATAATAAATAAATTAATCTTAAATGAAACGTCGATGAACGCACTGAAATAACACAGA 2545
Qy 747 CTATGATGTCAAATAAATAATCTTAAATGAAACGTCGATGAACGCACTGAAATAACACAGA 806
Db 2546 CTATGATGTCAAATAAATAATCTTAAATGAAACGTCGATGAACGCACTGAAATAACACAGA 2605
Qy 807 TATCAAAAGTTGTGACTTTGGAAGTGCAACGATGATGATGATGATGATGATGATGATGATGAT 866
Db 2606 TATCAAAAGTTGTGACTTTGGAAGTGCAACGATGATGATGATGATGATGATGATGATGATGAT 2665
Qy 867 GTCTACCGGCACTACAGACTCCGAGGTCATTTTGGCTTTAGTGGTCTCAGCCTTG 926
Db 2666 GTCTACCGGCACTACAGACTCCGAGGTCATTTTGGCTTTAGTGGTCTCAGCCTTG 2725
Qy 927 TGATGTTTGGAGCATAGGTTGCATTTTATTAATGAAATATTACCTTGGTTTCAACAGTCTTTCA 986
Db 2726 TGATGTTTGGAGCATAGGTTGCATTTTATTAATGAAATATTACCTTGGTTTCAACAGTCTTTCA 2785
Qy 987 GACTCATGATGATAAGAGCACCTGGCAATGATGGAACGAATATPAGGACCCATACCA 1046
Db 2786 GACTCATGATGATAAGAGCACCTGGCAATGATGGAACGAATATPAGGACCCATACCA 2845
Qy 1047 ACACATGATTCAGAAAAACAGAAAAACGCAAGTATTTTCCACATACAGCTAGATTGGGA 1106
Db 2846 ACACATGATTCAGAAAAACAGAAAAACGCAAGTATTTTCCACATACAGCTAGATTGGGA 2905
Qy 1107 TGAACACAGTTTCTGCTGTTGATATGTTTAGGAGACGCTGCAACCGTTGAAAGGAATTTAT 1166
Db 2906 TGAACACAGTTTCTGCTGTTGATATGTTTAGGAGACGCTGCAACCGTTGAAAGGAATTTAT 2965
Qy 1167 GCTTTGTCATGATGAAGAACATGAGAAAACGTTTGACCTGGTTCGAAGAATGTTAGAATA 1226
Db 2966 GCTTTGTCATGATGAAGAACATGAGAAAACGTTTGACCTGGTTCGAAGAATGTTAGAATA 3025
Qy 1227 TGATCCAACTCAAGAAATTAACCTTGGATGAAGCATTCAGCATCTCTTTCTTTGACTTAT 1286
Db 3026 TGATCCAACTCAAGAAATTAACCTTGGATGAAGCATTCAGCATCTCTTTCTTTGACTTAT 3085
Qy 1287 AAAAAAGAAATGAATGGAAATCAGTGGTCTTACTATATATCTCTCTAGAGAGATTA 1346
Db 3086 AAAAAAGAAATGAATGGAAATCAGTGGTCTTACTATATATCTCTCTAGAGAGATTA 3145
Qy 1347 TAAGACTGTGTCAGTCAACTAAACATCTTAATATTTTGTAAACATTAATATTTTGTGA 1406
Db 3146 TAAGACTGTGTCAGTCAACTAAACATCTTAATATTTTGTAAACATTAATATTTTGTGA 3205
Qy 1407 CAGTTAAGTGAATAATTTGATGTTTGTATCAATPAGCATTAATTAATCTTTGTAAGCAAGT 1466
Db 3206 CAGTTAAGTGAATAATTTGATGTTTGTATCAATPAGCATTAATTAATCTTTGTAAGCAAGT 3265
Qy 1467 ATGGTCTTTGATAATGCATTAAGAAAATTAATAATTTTCTTTTTCGAAATTAACATTT 1526
Db 3266 ATGGTCTTTGATAATGCATTAAGAAAATTAATAATTTTCTTTTTCGAAATTAACATTT 3325
Qy 1527 TTAATACTCTTTGAAATATCTTTGTCGAGTGAATAATGATGATCTTTGCTCTTTG 1586
Db 3326 TTAATACTCTTTGAAATATCTTTGTCGAGTGAATAATGATGATCTTTGCTCTTTG 3385
Qy 1587 TACATGGAGGTCACTCTGAAATGATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 1646
```

```
Db 3386 TACATGGAGGTCACTCTGAAAGTGAATTTTTTTTGAAGTAAAGGAAATCTTGACTACTTTA 3445
Qy 1647 TATTTCTTAAAGGAATATTTCTTTTATATCTCAAAATTTAGAACTTAACCTTTTAAAGTTTTT 1706
Db 3446 TATTTCTTAAAGGAATATTTCTTTTATATCTCAAAATTTAGAACTTAACCTTTTAAAGTTTTT 3505
Qy 1707 CTTCTGTAATTTGTTGAACGGGTGATTTATTAATCTAGATAAGCAGGTACTAGAAACC 1766
Db 3506 CTTCTGTAATTTGTTGAACGGGTGATTTATTAATCTAGATAAGCAGGTACTAGAAACC 3565
Qy 1767 AAAAATCAGAAAATGTTTACTGTTAGTAATCTTATTAATTTTAAAGTGTGTTGTTCTTTT 1826
Db 3566 AAAAATCAGAAAATGTTTACTGTTAGTAATCTTATTAATTTTAAAGTGTGTTGTTCTTTT 3625
Qy 1827 CATTTGGGTGATGTGAGGTGATAACCCAGACATTCATGAAAGGCCATGCGATTTGTCATTT 1886
Db 3626 CATTTGGGTGATGTGAGGTGATAACCCAGACATTCATGAAAGGCCATGCGATTTGTCATTT 3685
Qy 1887 GTGACAGTTTGTGTTTAAATAAAACCCACATACACATTTTATTAAGATTAAATCTTAACCTGA 1946
Db 3686 GTGACAGTTTGTGTTTAAATAAAACCCACATACACATTTTATTAAGATTAAATCTTAACCTGA 3745
Qy 1947 AAGTCAGCTTCGAAAATGCGACATTTTCCAAGTATGTTTGGTGTGAGTCCACAGATATAAAATA 2006
Db 3746 AAGTCAGCTTCGAAAATGCGACATTTTCCAAGTATGTTTGGTGTGAGTCCACAGATATAAAATA 3805
Qy 2007 GAAATTTCTGATGAGAGGTTTTCAGTTTTTAAATCAACAGTCTTTAGGAGTCTTAAACATTTGCG 2066
Db 3806 GAAATTTCTGATGAGAGGTTTTCAGTTTTTAAATCAACAGTCTTTAGGAGTCTTAAACATTTGCG 3865
Qy 2067 CAGCATCTGTTTATCAAAATGACATAAATACGTAAACCTATTAAGAAATTAAGTTTATTAAT 2126
Db 3866 CAGCATCTGTTTATCAAAATGACATAAATACGTAAACCTATTAAGAAATTAAGTTTATTAAT 3925
Qy 2127 AGGCAATTTATGCTGTGATTAATTTCTTACGGGAGAAAGAGATTTGATTGGAAGCAGTT 2186
Db 3926 AGGCAATTTATGCTGTGATTAATTTCTTACGGGAGAAAGAGATTTGATTGGAAGCAGTT 3985
Qy 2187 TGGGAAGAAAGTGTGCTGTAATTTTCCAAGATTTTAAATGATTGTTTACAT 2236
Db 3986 TGGGAAGAAAGTGTGCTGTAATTTTCCAAGATTTTAAATGATTGTTTACAT 4035
```

RESULT 6

```
US-10-267-502-135
; Sequence 135: Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 135
; LENGTH: 1446
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-267-502-135
```

```
Query Match 57.0%; Score 1286; DB 17; Length 1446;
Best Local Similarity 100.0%; Pred. No. 2.8e-260;
Matches 1286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 GTCATTTTAAAGCAAGGTCCTTGAATGACGAGATTAATCGGACCGGAGATACGTTG 73
Db 161 GTCATTTTAAAGCAAGGTCCTTGAATGACGAGATTAATCGGACCGGAGATACGTTG 220
Qy 74 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTTATCACAGACATTTG 133
Db 221 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTTATCACAGACATTTG 280
```

Qy	134	AAAGCGGGTATCGAATCCATCGCAGTAAATCTTCAGTCCGACGAGGAAGACGAGTCTTA	193
Db	281	AAAGCGGGTATCGAATCCATCGCAGTAAATCTTCAGTCCGACGAGGAAGACGAGTCTTA	340
Qy	194	AAAGGAAGCGCAATAGACACATGTTCAAGTCATCATGTCACGTTCCGAGAGCCACCGAAGGA	253
Db	341	AAAGGAAGCGCAATAGACACATGTTCAAGTCATCATGTCACGTTCCGAGAGCCACCGAAGGA	400
Qy	254	AAAGATCCAGGAGTATAGAGATGATGAGGAGGGTCACCTGATCTGTCAAAAGTCGAGACG	313
Db	401	AAAGATCCAGGAGTATAGAGATGATGAGGAGGGTCACCTGATCTGTCAAAAGTCGAGACG	460
Qy	314	TTCTAAGAGCAAGATATGAATCTGTGGACACTTTTGGGTGAAGAGCCCTTTGGCAAAAGTTG	373
Db	461	TTCTAAGAGCAAGATATGAATCTGTGGACACTTTTGGGTGAAGAGCCCTTTGGCAAAAGTTG	520
Qy	374	TAGAGTGCATTTGATCATGGCATGATGCGATGTCATGTAGCAGTCGAAATCGTAAAAAATG	433
Db	521	TAGAGTGCATTTGATCATGGCATGATGCGATGTCATGTAGCAGTCGAAATCGTAAAAAATG	580
Qy	434	TAGCCCGTTACCGTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTA	493
Db	581	TAGCCCGTTACCGTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTA	640
Qy	494	CTGATCCCAATAGTGTCTTCGATGTGTCCAGATGCTAGAAATGTTTGATCATCATGGTC	553
Db	641	CTGATCCCAATAGTGTCTTCGATGTGTCCAGATGCTAGAAATGTTTGATCATCATGGTC	700
Qy	554	ATGTTTGTATTGTGTTTGAACCTACTGGGACCTTAGTACTTACGATTTCAATTAAGAAAAACA	613
Db	701	ATGTTTGTATTGTGTTTGAACCTACTGGGACCTTAGTACTTACGATTTCAATTAAGAAAAACA	760
Qy	614	GCTTTCGCCATTTCAAATTTGACCAACATCAGGCGAGATGGCGTATCAGATCTGCCAGTCAA	673
Db	761	GCTTTCGCCATTTCAAATTTGACCAACATCAGGCGAGATGGCGTATCAGATCTGCCAGTCAA	820
Qy	674	TAAATTTTTTACATCATATAAATTAACCCATACAGATCTCAAGCTCGAAATATTTTGT	733
Db	821	TAAATTTTTTACATCATATAAATTAACCCATACAGATCTCAAGCTCGAAATATTTTGT	880
Qy	734	TTGTGAAGCTTGACATGTAGTCAAAATATAATTTCAAATGAAACGTCGATGAACGCACAC	793
Db	881	TTGTGAAGCTTGACATGTAGTCAAAATATAATTTCAAATGAAACGTCGATGAACGCACAC	940
Qy	794	TGAAAAACACAGATATCAAGTGTGTGACCTTTGGAAGTCGAACTGATCATGAACATC	853
Db	941	TGAAAAACACAGATATCAAGTGTGTGACCTTTGGAAGTCGAACTGATCATGAACATC	1000
Qy	854	ACAGTACTTTGGTGTCTACCCGGCAGCTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGGTT	913
Db	1001	ACAGTACTTTGGTGTCTACCCGGCAGCTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGGTT	1060
Qy	914	GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCAATTCCTTATTGAAATATTACCTTGGTT	973
Db	1061	GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCAATTCCTTATTGAAATATTACCTTGGTT	1120
Qy	974	TCACAGCTTTTCAGACTCATGATAGTAAAGAGCAGCTGGCAATGATGGAACGAAATATTAG	1033
Db	1121	TCACAGCTTTTCAGACTCATGATAGTAAAGAGCAGCTGGCAATGATGGAACGAAATATTAG	1180
Qy	1034	GACCCATACCAACAACATGATTAGAAAAACAAGAAACGCAAGTATTTTCCACCATACCC	1093
Db	1181	GACCCATACCAACAACATGATTAGAAAAACAAGAAACGCAAGTATTTTCCACCATACCC	1240
Qy	1094	AGCTTAGATTGGGATGAACACAGTTCTGCTGCTAGATATGTTAGGAGAGCTTGCAACCGT	1153
Db	1241	AGCTTAGATTGGGATGAACACAGTTCTGCTGCTAGATATGTTAGGAGAGCTTGCAACCGT	1300
Qy	1154	TGAAGGAATTTTATGCTTTTGTTCATGATGAAGAAACATGAAAACTGTTTGACCTGGTTCCAA	1213
Db	1301	TGAAGGAATTTTATGCTTTTGTTCATGATGAAGAAACATGAAAACTGTTTGACCTGGTTCCAA	1360

Qy	1214	GAATGTTAGAAATATGATCCAACTCAAGAATAATTACCTTGGATGAAGCATTCGACATCCTTT	1273
Dd	1361	GAATGTTAGAAATATGATCCAACTCAAGAATAATTACCTTGGATGAAGCATTCGACATCCTTT	
Qy	1274	TCCTTGACTATTAAAAAAGAAAATGA	1299
Dd	1421	TCCTTGACTATTAAAAAAGAAAATGA	1446
 RESULT 7 US-10-641-643-699 ; Sequence 699, Application US/10641643 ; Publication No. US20040077003A1 ; GENERAL INFORMATION: ; APPLICANT: Cocks, Benjamin G. ; Susan G. Stuart ; Jeffrey J. Seilhamer ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL ; GENE EXPRESSION ; NUMBER OF SEQUENCES: 1508 ; CORRESPONDENCE ADDRESS: ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC. ; STREET: 3174 PORTER DRIVE ; CITY: PALO ALTO ; STATE: CALIFORNIA ; COUNTRY: USA ; ZIP: 94304 ; COMPUTER READABLE FORM: ; MEDIUM TYPE: Floppy disk ; COMPUTER: IBM PC compatible ; OPERATING SYSTEM: PC-DOS/MS-DOS ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2 ; CURRENT APPLICATION DATA: ; APPLICATION NUMBER: US/10/641,643 ; FILING DATE: 14-Aug-2003 ; CLASSIFICATION: <Unknown> ; PRIOR APPLICATION DATA: ; APPLICATION NUMBER: <Unknown> ; FILING DATE: <Unknown> ; ATTORNEY/AGENT INFORMATION: ; NAME: Zeller, Karen J. ; REGISTRATION NUMBER: 37,071 ; REFERENCE/DOCKET NUMBER: PA-0001 US ; TELECOMMUNICATION INFORMATION: ; TELEPHONE: (650) 855-0555 ; TELEFAX: (650) 845-4166 ; INFORMATION FOR SEQ ID NO: 699: ; SEQUENCE CHARACTERISTICS: ; LENGTH: 1456 base pairs ; TYPE: nucleic acid ; STRANDEDNESS: single ; TOPOLOGY: linear ; IMMEDIATE SOURCE: ; LIBRARY: HNT2AGT01 ; CLONE: 488842 ; SEQUENCE DESCRIPTION: SEQ ID NO: 699 : US-10-641-643-699			
 Query Match 52.2%; Score 1178.2; DB 17; Length 1456; Best Local Similarity 99.7%; Pred. No. 1.4e-237; Matches 1191; Conservative 0; Mismatches 3; Indels 1; Gaps 1;			
Qy	932	TTTTGGACATAGGTGGCATCTTATTGAATATTACCTTGGTTTTACAGTCTTTTCAGACTC	991
Dd	262	TGTAGCGCATAGGTTGCATCTTATTGAATATTACCTTGGTTTTACAGTCTTTTCAGACTC	321
Qy	992	ATGATAGTAAGAAGACA CTGGCAATGATGGAACGAATATTAGGACCCTATCCCAACACA	1051
Dd	322	ATGATAGTAAGAAGCACCTGGCAATGATGGAACGAATATTAGGACCCTATCCCAACACA	381
Qy	1052	TGATTCAGAAAAACAAGAAAACGGAAGTATTTTTCACATACCAAGTAGATTGGATGAAC	1111
Dd	382	TGATTCAGAAAAACAAGAAAACGGAAGTATTTTTCACATACCAAGTAGATTGGATGAAC	441

Db 1259 GACCCATCCAGCACATATGATCCAGAGACCAAGGAACCGCAAGTATTTCCACCAATAACC 1318
QY 1094 AGCTAGATTGGGATGAACACAGCTTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGT 1153
Db 1319 AGCTAGATTGGGACGAGCATAGTTCACTGGGAGATATGTTAGGAGACGCTGCAAGCCGT 1378
QY 1154 TGAAGAAATTTATGCTTTGCTATGATGAAGAAACATGAGAAACATGTTGACCTGGTTGGAA 1213
Db 1379 TAAAGAAATTTATGCTGCTATGATGACCAAGAGCATGAGAAGCTGTTGACCTGGTTGGAA 1438
QY 1214 GAATGTTAGAAATGATCCAACTCAAGAAATTTACCTTGGATCAAGCATTTGAGGATCCCTT 1273
Db 1439 GAATGTTGAGATGATGACCCAGGAGAGGATCACCTTGGATGAAGCATTTGAGGATCCCTT 1498
QY 1274 TCTTTGACTTATTAAGAAAGAAATGAATGG 1304
Db 1499 TCTTTGACTTATTAAGAAAGAAATGATGTTGG 1529

RESULT 10

US-10-267-502-138
; Sequence 138, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 138
; LENGTH: 1446
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-267-502-138

Query Match 48.6%; Score 1097.2; DB 17; Length 1446;
Best Local Similarity 90.8%; Pred. No. 1.5e-220;
Matches 1168; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 14 GTCACTATTAGAAGCAAGCTCTTGAATGACGAGATTTATCGGACCGGAGATACGTTG 73
Db 161 GTCACTATTAGAAGCAAGATCTTGAATGAGAGATTTATCGGACCGGAGATACATTG 220
QY 74 AGCAATACAGGAATGACTACTGTGAAGGATATGTTCTAGACATTTATCAAGAGACATTG 133
Db 221 ATGAATACAGAAATGACTACTCGGAAGGATATGTTCTCAAGACATTTACCATAGAGACGTTG 280
QY 134 AAAGCGGTATCGAATCCACTCGAGTAAATCTTCAGTCCGACGAGGAGAGCAGTCCTTA 193
Db 281 AAAGCACTTACCGGATCCATTGAGTAAATCTTCAGTCCGACGAGGAGAGCAGTCCTTA 340
QY 194 AAAGGAAGCGCAATAGACACTGTTCAAGTCACTAGTCAGTTCGAGAGGCCACCGAAGGA 253
Db 341 AGAGAAAGCGTAAATAGACCTCTGCAAGTCACTAGTCAGTTCGAGAGGCCACCGAAGGA 400
QY 254 AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGTCMAAGTGAGAGG 313
Db 401 AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGTCMAAGTGAGAGG 460
QY 314 TTCTAAGAGCAAGATATGAATCGTGGACACATTTGGTGAAGGAGCCTTTGCAAGTTG 373
Db 461 TTCTAAGAGCAAGATATGAATCGTGGACACATTTAGTGAAGGAGCCTTTGCAAGTTG 520
QY 374 TAGAGTGCAATTGATCATGCGCATGGGATGCGATGATGAGCAGTGAAGAAATCGTAAAGGATG 433
Db 521 TAGAGTGCAATTGATCATGCGCATGGGATGCGTACATGATGAGCAGTGAAGAAATCGTAAAGGATG 580
QY 434 TAGGCGGTTACCGTGAAGAGCAGTCTGTTCAAGAAATCCAAAGTATTTAGAGCAGCTTAAATAGTA 493

Db 581 TAGGAGCTTACCGGAGGAGCAGCTCTTCTGAAATCCAAATATTGGAGCACTTTGAACAGCA 640
QY 494 CTGATCCCAATAGTCTCTTCCGATGTGTCAGATGTCAGAAATGGTTTGAATCATCATGTC 553
Db 641 CTGACCCCAACAGTCTCTTCCGATGTCGTCAGATGTCAGAGTGGTTTGAATCATCATGTC 700
QY 554 ATGTTTGTATTGTTGTTTGAACCTA CTGGGACCTTAGTACTTACGATTTCAATTAAGAAAAACA 613
Db 701 ATGTTTGTATTGTTGTTTGAACCTA CTGGGACCTTAGTACTTACGATTTCAATTAAGAAAAACA 760
QY 614 GCTTTCTGCCATTTCAAAATTCACCATCAGCAGATGCGGTATCAGATCTGCCAGTCAA 673
Db 761 GTTTCTGCCATTTCAAAATTCACCATCAGCAGATGCGGTATCAGATCTGCCAGTCAA 820
QY 674 TAAATTTTTTACATCATATAATTAACCCATACAGATCTGAAGCTGGAATAATTTTGT 733
Db 821 TAAATTTTTTACATCATATAATTAACCCATACAGATCTGAAGCTGGAATAATTTTAT 880
QY 734 TTGTGAAGTCTGACTATGATGTAATAATAATTAATTAATAAAGAAACGTAAGAACGACAC 793
Db 881 TTGTGAAGTCTGACTATGATGTAATAATAATTAATTAATAAAGAAACGTAAGAACGACAC 940
QY 794 TGAAGAAACACAGATATCAAGTGTGTTGACTTTGGAGTGCACGATGATGATGATGATGATG 853
Db 941 TGAAGAAACACAGATATCAAGTGTGTTGACTTTGGAGTGCACGATGATGATGATGATGATG 1000
QY 854 ACAGTACTTTGCTGTACTACCGGCACTACAGAGCTCCGAGGTCAATTTTGGCTTTAGGTT 913
Db 1001 ATAGTACTTTGCTGTACTACCGGCACTACAGAGGTCAATTTTGGCTTTAGGTT 1060
QY 914 GGTCTCAGCCTTTGATGATTTGGAGCATAGGTTGCAATTTTATTAATTAATTAATTAATTAAT 973
Db 1061 GGTCTCAGCCTTTGATGATTTGGAGCATAGGTTGCAATTTTATTAATTAATTAATTAATTAAT 1120
QY 974 TCACAGTCTTTGAGTCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1033
Db 1121 TCACAGTCTTTGAGTCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1180
QY 1034 GACCCATACCAACACACATGATTTCAAGAAACCAAGAAACCGCAAGTATTTTCCACCAATAACC 1093
Db 1181 GACCCATACCAACACATGATTTCAAGAAACCAAGAAACCGCAAGTATTTTCCACCAATAACC 1240
QY 1094 AGCTAGATTGGGATGAACACAGTCTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGT 1153
Db 1241 AGCTAGATTGGGATGAACACAGTCTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGT 1300
QY 1154 TGAAGAAATTTATGCTTTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1213
Db 1301 TAAAGGAATTTTATGCTGTTGCTATGATGATGATGATGATGATGATGATGATGATGATGATG 1360
QY 1214 GAATGTTAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1273
Db 1361 GAATGTTAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1420
QY 1274 TCTTTGACTTATTAAGAAAGAAATGA 1299
Db 1421 TCTTTGACTTATTAAGAAAGAAATGA 1446

RESULT 11

US-09-810-671-3
; Sequence 3, Application US/09810671
; Publication No. US20020076783A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: THEREOF
; FILE REFERENCE: CL000758
; CURRENT APPLICATION NUMBER: US/09/810,671
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3									
; LENGTH: 21234									
; TYPE: DNA									
; ORGANISM: Human									
US-09-810-671-3									
Query Match									
Best Local Similarity 48.6%; Score 1097; DB 9; Length 21234;									
Matches 1097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1157	AGGAATTTATGCTTGTGATGAGAAACATGAGAACTGTTGACCTGGTTCGAAGAA	1216						
DB	18138	AGGAATTTATGCTTGTGATGAGAAACATGAGAACTGTTGACCTGGTTCGAAGAA	18197						
QY	1217	TGTTAGATATGATCCCACTCAAGAAATTTACCTGGATGAAGCAATGCAGCATCTTCT	1276						
DB	18198	TGTTAGATATGATCCCACTCAAGAAATTTACCTGGATGAAGCAATGCAGCATCTTCT	18257						
QY	1277	TTGACTTATTAAGAAATGAAATGGAATCAAGTGTCTTACTATATCTTCTCTAGA	1336						
DB	18258	TTGACTTATTAAGAAATGAAATGGAATCAAGTGTCTTACTATATCTTCTCTAGA	18317						
QY	1337	AGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTA	1396						
DB	18318	AGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTA	18377						
QY	1397	TTATTTTGTACAGTTAACTGTAAATATTGTATGTTTGTATCAATAGCATATTAACCTG	1456						
DB	18378	TTATTTTGTACAGTTAACTGTAAATATTGTATGTTTGTATCAATAGCATATTAACCTG	18437						
QY	1457	TTAAGCAAGTATGCTTGTGATGATGAAATGAAATGGAATGGAATGGAATGGAATGGA	1516						
DB	18438	TTAAGCAAGTATGCTTGTGATGATGAAATGAAATGGAATGGAATGGAATGGAATGGA	18497						
QY	1517	ATTACCATTTTAAATACCTTTGAAATATCCCTTGTCCAGTCATTAATGTCGATTCAT	1576						
DB	18498	ATTACCATTTTAAATACCTTTGAAATATCCCTTGTCCAGTCATTAATGTCGATTCAT	18557						
QY	1577	TTGCTTTTGTACATGAGGTCACCTCTGAAGTCATTTTTCAGTAAAGGAAATCTT	1636						
DB	18558	TTGCTTTTGTACATGAGGTCACCTCTGAAGTCATTTTTCAGTAAAGGAAATCTT	18617						
RESULT 12									
US-10-109-854-3									
; Sequence 3, Application US/10109854									
; Publication No. US20020119548A1									
; GENERAL INFORMATION:									
; APPLICANT: YAN, Chunhua et al.									
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC									
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES									
; TITLE OF INVENTION: THEREOF									
; FILE REFERENCE: CLO000758DIV									
; CURRENT APPLICATION NUMBER: US/10/109,854									
; CURRENT FILING DATE: 2002-04-01									
; PRIOR APPLICATION NUMBER: 60/227,470									
; PRIOR FILING DATE: 2000-08-24									
; PRIOR APPLICATION NUMBER: 09/810,671									
; PRIOR FILING DATE: 2001-03-19									
; NUMBER OF SEQ ID NOS: 5									
; SOFTWARE: Fast-SEQ for Windows Version 4.0									
; SEQ ID NO 3									
; LENGTH: 21234									
; TYPE: DNA									
; ORGANISM: Homo sapien									
US-10-109-854-3									
Query Match									
Best Local Similarity 48.6%; Score 1097; DB 13; Length 21234;									
Matches 1097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1157	AGGAATTTATGCTTGTGATGAGAAACATGAGAACTGTTGACCTGGTTCGAAGAA	1216						
DB	18138	AGGAATTTATGCTTGTGATGAGAAACATGAGAACTGTTGACCTGGTTCGAAGAA	18197						
QY	1217	TGTTAGATATGATCCCACTCAAGAAATTTACCTGGATGAAGCAATGCAGCATCTTCT	1276						
DB	18198	TGTTAGATATGATCCCACTCAAGAAATTTACCTGGATGAAGCAATGCAGCATCTTCT	18257						
QY	1277	TTGACTTATTAAGAAATGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGA	1336						
DB	18258	TTGACTTATTAAGAAATGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGA	18317						
QY	1337	AGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTA	1396						
DB	18318	AGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTA	18377						
QY	1397	TTATTTTGTACAGTTAACTGTAAATATTGTATGTTTGTATCAATAGCATATTAACCTG	1456						
DB	18378	TTATTTTGTACAGTTAACTGTAAATATTGTATGTTTGTATCAATAGCATATTAACCTG	18437						
QY	1457	TTAAGCAAGTATGCTTGTGATGATGAAATGAAATGGAATGGAATGGAATGGAATGGA	1516						
DB	18438	TTAAGCAAGTATGCTTGTGATGATGAAATGAAATGGAATGGAATGGAATGGAATGGA	18497						
QY	1517	ATTACCATTTTAAATACCTTTGAAATATCCCTTGTCCAGTCATTAATGTCGATTCAT	1576						
DB	18498	ATTACCATTTTAAATACCTTTGAAATATCCCTTGTCCAGTCATTAATGTCGATTCAT	18557						
QY	1577	TTGCTTTTGTACATGAGGTCACCTCTGAAGTCATTTTTCAGTAAAGGAAATCTT	1636						
DB	18558	TTGCTTTTGTACATGAGGTCACCTCTGAAGTCATTTTTCAGTAAAGGAAATCTT	18617						
QY	1637	GACTACTTTATATCTTAAAGAAATATCTTTTATATATCTTCAAAATTTTGAACCTTA	1696						
DB	18618	GACTACTTTATATCTTAAAGAAATATCTTTTATATATCTTCAAAATTTTGAACCTTA	18677						
QY	1697	AAAGTTTTTCTTCTGTAATTTGTGAAACATGATGATGATGATGATGATGATGATG	1756						
DB	18678	AAAGTTTTTCTTCTGTAATTTGTGAAACATGATGATGATGATGATGATGATGATG	18737						
QY	1757	ACTAGAACCAAACTCAGAAATGCTTTTACCTGTTAGAAATCTTATTAATTTTAAAGTGTG	1816						
DB	18738	ACTAGAACCAAACTCAGAAATGCTTTTACCTGTTAGAAATCTTATTAATTTTAAAGTGTG	18797						
QY	1817	TATTTCTTTTCAATGGGTGATGTCAGGGTGATTAACACAGATTCATGAAAGGCGATGCG	1876						
DB	18798	TATTTCTTTTCAATGGGTGATGTCAGGGTGATTAACACAGATTCATGAAAGGCGATGCG	18857						
QY	1877	TTTGTCCATTTGACAGTTTGTGTTAAATAAACACACATACACATTTTATTAAGATTAAA	1936						
DB	18858	TTTGTCCATTTGACAGTTTGTGTTAAATAAACACACATACACATTTTATTAAGATTAAA	18917						
QY	1937	TCTAACTGGAAGTACAGTTTGTGAAATGGAATGGAATGGAATGGAATGGAATGGA	1996						
DB	18918	TCTAACTGGAAGTACAGTTTGTGAAATGGAATGGAATGGAATGGAATGGAATGGA	18977						
QY	1997	TATAAAATAGAAATTTCTGATGAGAGTTTCAAGTTTAAATACCAAGTCTTTAGAGTCT	2056						
DB	18978	TATAAAATAGAAATTTCTGATGAGAGTTTCAAGTTTAAATACCAAGTCTTTAGAGTCT	19037						
QY	2057	TAAACATTTGCCAGCATCTGTTTATCAATGACATTAATACCTTAAGATTAAAG	2116						
DB	19038	TAAACATTTGCCAGCATCTGTTTATCAATGACATTAATACCTTAAGATTAAAG	19097						

Db 19158 GAAAGCAGTTTGGGAGAAAGTGTCTGCTGAAATTTCCAGAAATTTAAATTGATGGTTACAT 19217
 QY 2237 AAACCTTTTGGACTTCAG 2253
 Db 19218 AAACCTTTTGGACTTCAG 19234

RESULT 14
 US-10-801-671-3
 ; Sequence 3, Application US/10801671
 ; Publication No. US20040152123A1
 ; GENERAL INFORMATION:
 ; APPLICANT: YAN, Chunhua et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: CL000758DIV-III
 ; CURRENT APPLICATION NUMBER: US/10/801,671
 ; CURRENT FILING DATE: 2004-03-17
 ; PRIOR APPLICATION NUMBER: 60/227,470
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: 09/810,671
 ; PRIOR FILING DATE: 2001-03-19
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 21234
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-801-671-3

Query Match 48.6%; Score 1097; DB 18; Length 21234;
 Best Local Similarity 100.0%; Pred. No. 6e-220;
 Matches 1097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1157 AGGAATTTATGCTTTGTCATGATGAAGAACATGAGAAACTGTGTCACCTGGTTCGAAGAA 1216
 Db 18138 AGGAATTTATGCTTTGTCATGATGAAGAACATGAGAAACTGTGTCACCTGGTTCGAAGAA 18197

QY 1217 TGTGATATATGATCCCACTCAAGAAATACCTTGGATGAAGCAATTCGAGCATCCTTTCT 1276
 Db 18198 TGTGATATATGATCCCACTCAAGAAATACCTTGGATGAAGCAATTCGAGCATCCTTTCT 18257

QY 1277 TTGACTTTATTAAGAAAGAAATGAAATGGAAATCAGTGTCTTACTATATATCTCTAGA 1336
 Db 18258 TTGACTTTATTAAGAAAGAAATGAAATGGAAATCAGTGTCTTACTATATATCTCTAGA 18317

QY 1337 AGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTCCTAATATTTTGTAAACATTTAA 1396
 Db 18318 AGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTCCTAATATTTTGTAAACATTTAA 18377

QY 1397 TTATTTTGTACAGTTAAGTGAATATTTCTATGTTTGTATCAATAGCATATTAACCTTG 1456
 Db 18378 TTATTTTGTACAGTTAAGTGAATATTTCTATGTTTGTATCAATAGCATATTAACCTTG 18437

QY 1457 TTAAGCAAGTATGGTCTTGATATGCAATGATAGAAATTAATAATTTTCTTTTGA 1516
 Db 18438 TTAAGCAAGTATGGTCTTGATATGCAATGATAGAAATTAATAATTTTCTTTTGA 18497

QY 1517 ATTACATTTTAAATACCTTTTGAATATTCCTTGTGTCAGTGAATATGATGATC 1576
 Db 18498 ATTACATTTTAAATACCTTTTGAATATTCCTTGTGTCAGTGAATATGATGATC 18557

QY 1577 TTGCTTTTGTACATGAGGTCACCTCTGAAGTGAATTTTGTGATGAAGAAATCTT 1636
 Db 18558 TTGCTTTTGTACATGAGGTCACCTCTGAAGTGAATTTTGTGATGAAGAAATCTT 18617

QY 1637 GACTACTTTATATCTTAAAGGAATATTTCTTATATATCTTCAAAATTTAGAACTTAACCTT 1696
 Db 18618 GACTACTTTATATCTTAAAGGAATATTTCTTATATATCTTCAAAATTTAGAACTTAACCTT 18677

QY 1697 AAAAGTTTCTTCTGTAATTTGAAACGGGTGATTTATTTAACTCTAGATAAGCAGT 1756

Db 18678 AAAAGTTTCTTCTGTAATTTGTAACGGGTGATTTATTTAACTCTAGATAAGCAGT 18737
 QY 1757 ACTAGAAACCAAAACTCAGAAATGTTTACTGTGTAGAAATCTTATTAATTTTAAAGTGTG 1816
 Db 18738 ACTAGAAACCAAAACTCAGAAATGTTTACTGTGTAGAAATCTTATTAATTTTAAAGTGTG 18797
 QY 1817 TATTCTTTTCAATTTGGGTGATGTCAGGTGATATACAGACATTCATATGAAAGGATGAG 1876
 Db 18798 TATTCTTTTCAATTTGGGTGATGTCAGGTGATATACAGACATTCATATGAAAGGATGAG 18857
 QY 1877 TTTCTCCATTTGTGACAGTTTCTTAATAAACCACATACACACTTTATTTAAAGATTAAA 1936
 Db 18858 TTTCTCCATTTGTGACAGTTTCTTAATAAACCACATACACACTTTATTTAAAGATTAAA 18917
 QY 1937 TCTAACTGGAAGTCAGCTTCGAAAATGGAATTTCCAAAGTATGTTGGTGAGTCACAGA 1996
 Db 18918 TCTAACTGGAAGTCAGCTTCGAAAATGGAATTTCCAAAGTATGTTGGTGAGTCACAGA 18977
 QY 1997 TATAAATAGAAATTCGTATGAGAGGTTTCAGTTTTTAATACCAAGTCCTTAGAGTCT 2056
 Db 18978 TATAAATAGAAATTCGTATGAGAGGTTTCAGTTTTTAATACCAAGTCCTTAGAGTCT 19037
 QY 2057 TAACTATGGCCAGCATCTGTTTATCAATGACATATAACCTTAAAGAAATTAAG 2116
 Db 19038 TAACTATGGCCAGCATCTGTTTATCAATGACATATAACCTTAAAGAAATTAAG 19097
 QY 2117 TTTATTAATTAGGCAATTTATGTCGTGATATTTCTTACGGGAGAAAGAGGATTTGATTG 2176
 Db 19098 TTTATTAATTAGGCAATTTATGTCGTGATATTTCTTACGGGAGAAAGAGGATTTGATTG 19157
 QY 2177 GAAAGCAGTTTGGGAGAAAGTGTCTGCTGAAATTTCCAGAAATTTAAATTGATGGTTACAT 2236
 Db 19158 GAAAGCAGTTTGGGAGAAAGTGTCTGCTGAAATTTCCAGAAATTTAAATTGATGGTTACAT 19217
 QY 2237 AAACCTTTTGGACTTCAG 2253
 Db 19218 AAACCTTTTGGACTTCAG 19234

RESULT 15
 US-10-425-114-26212
 ; Sequence 26212, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaska, Jack E.
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 26212
 ; LENGTH: 3040
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: LIB4115-001-H8_FLI
 US-10-425-114-26212

Query Match 43.3%; Score 977.4; DB 17; Length 3040;
 Best Local Similarity 99.9%; Pred. No. 3.5e-195;
 Matches 978; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 327 ATATGAAATCGTGGACACTTTGGGTGAAGAGCCCTTGGCAAAAGTTGTAGAGTCATTGA 386
 Db 2062 AGATGAAATCGTGGACACTTTGGGTGAAGAGCCCTTGGCAAAAGTTGTAGAGTCATTGA 2121
 QY 387 TCATGGCATGGATGGCATGTCATGATGAGTGAATATCGTAAATATGTAAGCCGTTACCG 446

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2005, 08:20:29 ; Search time 8553.23 Seconds
(without alignments)
10767.492 Million cell updates/sec

Title: US-10-801-671-1_COPY_72_2327

Perfect score: 2256

Sequence: 1 gacactatccagtcattt.....aaacttttgacttcagaaa 2256

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_NA_Main:

1: /cgn2_6/ptodata/1/pna/PCTUS1 COMB.seq.*
2: /cgn2_6/ptodata/1/pna/PCTUS2 COMB.seq.*
3: /cgn2_6/ptodata/1/pna/PCTUS COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US06 COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US07 COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US081 COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US082 COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US083 COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US084 COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US085 COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US086 COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US087 COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US088 COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US089 COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US090 COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US091 COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US092A COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US092B COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US093A COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US093B COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US094 COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US095A COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US095B COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US095C COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US095D COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096A COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US096B COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US096C COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US096D COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US096E COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US097A COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US097B COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US097C COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US098A COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US098B COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US098C COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US098D COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US099A COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US099B COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US099C COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US099D COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US099E COMB.seq.*

44: /cgn2_6/ptodata/1/pna/US099F COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US099G COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US100A COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US100B COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US101A COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US101B COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US102A COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US102B COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US103A COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US103B COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US104A COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US104B COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US105A COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US105B COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US106A COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US107A COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US107B COMB.seq.*
61: /cgn2_6/ptodata/1/pna/US107C COMB.seq.*
62: /cgn2_6/ptodata/1/pna/US107D COMB.seq.*
63: /cgn2_6/ptodata/1/pna/US108A COMB.seq.*
64: /cgn2_6/ptodata/1/pna/US108B COMB.seq.*
65: /cgn2_6/ptodata/1/pna/US109A COMB.seq.*
66: /cgn2_6/ptodata/1/pna/US109B COMB.seq.*
67: /cgn2_6/ptodata/1/pna/US109C COMB.seq.*
68: /cgn2_6/ptodata/1/pna/US110 COMB.seq.*
69: /cgn2_6/ptodata/1/pna/US6000 COMB.seq.*
70: /cgn2_6/ptodata/1/pna/US6001 COMB.seq.*
71: /cgn2_6/ptodata/1/pna/US6002 COMB.seq.*
72: /cgn2_6/ptodata/1/pna/US6003 COMB.seq.*
73: /cgn2_6/ptodata/1/pna/US6004 COMB.seq.*
74: /cgn2_6/ptodata/1/pna/US6005 COMB.seq.*
75: /cgn2_6/ptodata/1/pna/US6006 COMB.seq.*
76: /cgn2_6/ptodata/1/pna/US6007 COMB.seq.*
77: /cgn2_6/ptodata/1/pna/US6008 COMB.seq.*
78: /cgn2_6/ptodata/1/pna/US6009 COMB.seq.*
79: /cgn2_6/ptodata/1/pna/US6010 COMB.seq.*
80: /cgn2_6/ptodata/1/pna/US6011 COMB.seq.*
81: /cgn2_6/ptodata/1/pna/US6012 COMB.seq.*
82: /cgn2_6/ptodata/1/pna/US6013 COMB.seq.*
83: /cgn2_6/ptodata/1/pna/US6014 COMB.seq.*
84: /cgn2_6/ptodata/1/pna/US6015 COMB.seq.*
85: /cgn2_6/ptodata/1/pna/US6016 COMB.seq.*
86: /cgn2_6/ptodata/1/pna/US6017 COMB.seq.*
87: /cgn2_6/ptodata/1/pna/US6018 COMB.seq.*
88: /cgn2_6/ptodata/1/pna/US6019 COMB.seq.*
89: /cgn2_6/ptodata/1/pna/US6020 COMB.seq.*
90: /cgn2_6/ptodata/1/pna/US6021 COMB.seq.*
91: /cgn2_6/ptodata/1/pna/US6022 COMB.seq.*
92: /cgn2_6/ptodata/1/pna/US6023A COMB.seq.*
93: /cgn2_6/ptodata/1/pna/US6023B COMB.seq.*
94: /cgn2_6/ptodata/1/pna/US6024 COMB.seq.*
95: /cgn2_6/ptodata/1/pna/US6025 COMB.seq.*
96: /cgn2_6/ptodata/1/pna/US6026 COMB.seq.*
97: /cgn2_6/ptodata/1/pna/US6027 COMB.seq.*
98: /cgn2_6/ptodata/1/pna/US6028 COMB.seq.*
99: /cgn2_6/ptodata/1/pna/US6029 COMB.seq.*
100: /cgn2_6/ptodata/1/pna/US6030 COMB.seq.*
101: /cgn2_6/ptodata/1/pna/US6031 COMB.seq.*
102: /cgn2_6/ptodata/1/pna/US6032 COMB.seq.*
103: /cgn2_6/ptodata/1/pna/US6033 COMB.seq.*
104: /cgn2_6/ptodata/1/pna/US6034 COMB.seq.*
105: /cgn2_6/ptodata/1/pna/US6035 COMB.seq.*
106: /cgn2_6/ptodata/1/pna/US6036 COMB.seq.*
107: /cgn2_6/ptodata/1/pna/US6037 COMB.seq.*
108: /cgn2_6/ptodata/1/pna/US6038 COMB.seq.*
109: /cgn2_6/ptodata/1/pna/US6039 COMB.seq.*
110: /cgn2_6/ptodata/1/pna/US6040 COMB.seq.*
111: /cgn2_6/ptodata/1/pna/US6041 COMB.seq.*
112: /cgn2_6/ptodata/1/pna/US6042 COMB.seq.*
113: /cgn2_6/ptodata/1/pna/US6043 COMB.seq.*
114: /cgn2_6/ptodata/1/pna/US6044 COMB.seq.*
115: /cgn2_6/ptodata/1/pna/US6045 COMB.seq.*
116: /cgn2_6/ptodata/1/pna/US6046 COMB.seq.*

117: /cgn2_6/prodata/1/pna/US6047_COMB.seq:
118: /cgn2_6/prodata/1/pna/US6048_COMB.seq:
119: /cgn2_6/prodata/1/pna/US6049_COMB.seq:
120: /cgn2_6/prodata/1/pna/US6050_COMB.seq:
121: /cgn2_6/prodata/1/pna/US6051_COMB.seq:
122: /cgn2_6/prodata/1/pna/US6052_COMB.seq:
123: /cgn2_6/prodata/1/pna/US6053_COMB.seq:
124: /cgn2_6/prodata/1/pna/US6054_COMB.seq:
125: /cgn2_6/prodata/1/pna/US6055_COMB.seq:
126: /cgn2_6/prodata/1/pna/US6056_COMB.seq:
127: /cgn2_6/prodata/1/pna/US6057_COMB.seq:
128: /cgn2_6/prodata/1/pna/US6058_COMB.seq:
129: /cgn2_6/prodata/1/pna/US6059_COMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
------------	-------	-------------	--------	----	----	-------------

1	2256	100.0	2354	63	US-10-801-671-1	Sequence 1, Appli
2	2215.8	98.2	2497	48	US-10-111-076A-1	Sequence 1, Appli
3	2048	90.8	2488	44	US-09-979-167-55	Sequence 55, Appl
4	2047	90.7	2443	102	US-60-324-185-20117	Sequence 20117, A
5	2047	90.7	2487	97	US-60-278-358-3721	Sequence 3721, Ap
6	2046.6	90.7	2446	21	US-09-332-978-9	Sequence 9, Appli
7	1908.4	84.6	4035	54	US-10-425-114-26852	Sequence 26852, A
8	1908.4	84.6	4035	54	US-10-425-114-26852	Sequence 26852, A
9	1825	80.9	2494	1	PCT-US01-14827-2154	Sequence 2154, Ap
10	1825	80.9	2494	26	US-09-577-408-3087	Sequence 3087, Ap
11	1635	72.5	1795	45	US-10-170-335-20526	Sequence 20526, A
12	1634.6	72.5	1795	65	US-10-932-349-580	Sequence 580, App
13	1634.6	72.5	1795	120	US-60-500-337-580	Sequence 580, App
14	1632	72.3	1814	62	US-10-756-149-49	Sequence 49, Appl
15	1632	72.3	1814	117	US-60-474-733-338	Sequence 338, App
16	1300.6	57.7	1323	43	US-09-959-164-7	Sequence 7, Appli
17	1300.6	57.7	1323	43	US-09-959-164-7	Sequence 7, Appli
18	1286	57.0	1446	51	US-10-267-502-135	Sequence 135, App
19	1263.2	56.0	1348	29	US-09-652-128-8650	Sequence 8650, Ap
20	1263.2	56.0	1348	29	US-09-652-128-8650	Sequence 8650, Ap
21	1263.2	56.0	1348	35	US-09-813-206-755	Sequence 755, App
22	1260	55.9	1261	90	US-60-213-178-1312	Sequence 1312, Ap
23	1228.4	54.5	1247	22	US-09-471-275-6193	Sequence 6193, Ap
24	1228.4	54.5	1247	22	US-09-471-275-6193	Sequence 6193, Ap
25	1228.4	54.5	1247	26	US-09-488-725B-5425	Sequence 5425, Ap
26	1185	52.5	1222	1	PCT-US00-06765-17	Sequence 17, Appl
27	1185	52.5	1222	1	PCT-US00-06765-17	Sequence 17, Appl
28	1185	52.5	1222	42	US-09-950-083-657	Sequence 657, App
29	1185	52.5	1222	42	US-09-950-083-657	Sequence 657, App
30	1185	52.5	1222	48	US-10-105-299-1157	Sequence 1157, Ap
31	1185	52.5	1222	64	US-10-868-184-657	Sequence 657, App
32	1185	52.5	1222	64	US-10-868-184A-657	Sequence 657, App
33	1185	52.5	1222	64	US-10-868-184B-657	Sequence 657, App
34	1185	52.5	1222	64	US-10-868-194B-657	Sequence 657, App
35	1185	52.5	1222	64	US-10-896-164-657	Sequence 657, App
36	1178.2	52.2	1456	58	US-10-641-643-699	Sequence 699, App
37	1099	48.7	1531	48	US-10-144-771-2809	Sequence 2809, Ap
38	1099	48.7	1531	106	US-60-360-207-2809	Sequence 2809, Ap
39	1097.4	48.6	1549	17	US-09-137-248C-26	Sequence 26, Appl
40	1097.4	48.6	1549	17	US-09-137-248C-26	Sequence 26, Appl
41	1097.4	48.6	1549	43	US-09-959-164A-12	Sequence 12, Appl
42	1097.4	48.6	1549	63	US-10-825-177-26	Sequence 26, Appl
43	1097.4	48.6	1549	117	US-60-474-733-130	Sequence 130, App
44	1097.2	48.6	1446	51	US-10-267-502-138	Sequence 138, App
45	1097	48.6	21234	63	US-10-801-671-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-10-801-671-1
; Sequence 1, Application US/10801671
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000758DIV-III
; CURRENT APPLICATION NUMBER: US/10/801,671
; PRIOR FILING DATE: 2004-03-17
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/810,671
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-801-671-1

Query Match
Best Local Similarity 100.0%; Score 2256; DB 63; Length 2354;
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACTCATCCAGTCATTATTAGAACGAGGTCCTTGAATGAGCGAGATTATCGGAC 60
DB 72 GACACTCATCCAGTCATTATTAGAACGAGGTCCTTGAATGAGCGAGATTATCGGAC 131
QY 61 CGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTAT 120
DB 132 CGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTAT 191
QY 121 CACGAGACATTTGAAAGCGGGTATCGAATCCACTGCAGTAATCTTCTAGTCGGACGAG 180
DB 192 CACGAGACATTTGAAAGCGGGTATCGAATCCACTGCAGTAATCTTCTAGTCGGACGAG 251
QY 181 AGAAGCAGTCTTAAAGGAGCGCAATAGACACTGTTTCAAGTCATCATCTACGTTCCGAAG 240
DB 252 AGAAGCAGTCTTAAAGGAGCGCAATAGACACTGTTTCAAGTCATCATCTACGTTCCGAAG 311
QY 241 AGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATGAGGAGGGTCACTGATCTGT 300
DB 312 AGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATGAGGAGGGTCACTGATCTGT 371
QY 301 CAAAGTGAGACGTTCTTAAGAGCAAGATATGAATCGTGGACACTTTGGGTGAAGAGCC 360
DB 372 CAAAGTGAGACGTTCTTAAGAGCAAGATATGAATCGTGGACACTTTGGGTGAAGAGCC 431
QY 361 TTTGCAAAAGTTGTAGAGTCATTGATCATGGCATGGATGGCATGTCATGTAGCAGTGA 420
DB 432 TTTGCAAAAGTTGTAGAGTCATTGATCATGGCATGGATGGCATGTCATGTAGCAGTGA 491
QY 421 ATCTGAAAAATGTAGGCGGTTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAG 480
DB 492 ATCTGAAAAATGTAGGCGGTTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAG 551
QY 481 CACTTAAATAGTACTGATCCCAATAGTGTCTTCCGATGTGTCAGATGCTAGATGTTT 540
DB 552 CACTTAAATAGTACTGATCCCAATAGTGTCTTCCGATGTGTCAGATGCTAGATGTTT 611
QY 541 GATCATCATGTCTATGTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 600
DB 612 GATCATCATGTCTATGTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 671
QY 601 ATTAAGAAAAACAGCTTTCTCCCAATTCATTAAGTACCAATGACATGAGCAGATGCGGTATCAG 660
DB 672 ATTAAGAAAAACAGCTTTCTCCCAATTCATTAAGTACCAATGACATGAGCAGATGCGGTATCAG 731
QY 661 ATCTGCCAGTCAATAAATTTTTTACATCAATAAATTAACCCATACAGATCTGAAGCCT 720
|||||

Db 732 ATCTGCCAGTCAATAAATTTTACATCATATAATAAATAAACCACATACAGATCTGAAGCCT 791
 QY 721 GAAATATATTTTGTGTAAGTCTGACTATGTAGTCAAAATATAAATCTTAAATGAAACGT 780
 Db 792 GAAATATATTTTGTGTAAGTCTGACTATGTAGTCAAAATATAAATCTTAAATGAAACGT 851
 QY 781 GATGAACGCACACTGAAACACACAGATATCAAAAGTTGTGACTTTGGAAGTGCACAGTAT 840
 Db 852 GATGAACGCACACTGAAACACACAGATATCAAAAGTTGTGACTTTGGAAGTGCACAGTAT 911
 QY 841 GATGATGAACATCAAGTACTTTGTGTCTACCCGCGACTACAGAGCTCCCGAGGTCAAT 900
 Db 912 GATGATGAACATCAAGTACTTTGTGTCTACCCGCGACTACAGAGCTCCCGAGGTCAAT 971
 QY 901 TTGGCTTTAGGTGGTCTCAGCTTTGTGATGTTGGAGCATAGGTTCATTTATTGAA 960
 Db 972 TTGGCTTTAGGTGGTCTCAGCTTTGTGATGTTGGAGCATAGGTTCATTTATTGAA 1031
 QY 961 TATTACCTTTGGTTTTCAGAGTCTTTTCAGACTCATGATAGTAAGAGACCTCGGCAATGATG 1020
 Db 1032 TATTACCTTTGGTTTTCAGAGTCTTTTCAGACTCATGATAGTAAGAGACCTCGGCAATGATG 1091
 QY 1021 GAAAGAAATATTAGGACCCATACCAACACATGATTTTCAGAAACCAAGAAACGCAAGTAT 1080
 Db 1092 GAAAGAAATATTAGGACCCATACCAACACATGATTTTCAGAAACCAAGAAACGCAAGTAT 1151
 QY 1081 TTTTCAACATAACAGTATGGGATGAAACACAGTCTTGTGTTAGATATGTTAGGAGA 1140
 Db 1152 TTTTCAACATAACAGTATGGGATGAAACACAGTCTTGTGTTAGATATGTTAGGAGA 1211
 QY 1141 CGCTGCAACCGTTGGAAGAAATTTATGTTTGTCTGATGATGAAGAAACATGAGAAACCTGTTT 1200
 Db 1212 CGCTGCAACCGTTGGAAGAAATTTATGTTTGTCTGATGATGAAGAAACATGAGAAACCTGTTT 1271
 QY 1201 GACCTGGTTTGAAGAAATTTAGAAATGATGATCAACTCAAGAAATTAACCTTGGATGAAGCA 1260
 Db 1272 GACCTGGTTTGAAGAAATTTAGAAATGATGATCAACTCAAGAAATTAACCTTGGATGAAGCA 1331
 QY 1261 TTGACAGTCTCTTTCTTTGACTTATTAAGAAAGAAATGAAATGGAATCAGTGGTCTTAC 1320
 Db 1332 TTGACAGTCTCTTTCTTTGACTTATTAAGAAAGAAATGAAATGGAATCAGTGGTCTTAC 1391
 QY 1321 TATATACCTCTCTAGAGAGATTAAGTCTGTCAGTCACTAAACATCTTAATAT 1380
 Db 1392 TATATACCTCTCTAGAGAGATTAAGTCTGTCAGTCACTAAACATCTTAATAT 1451
 QY 1381 TTTTGTAAACATTAATATTTTGTACAGTTAAAGTAAATATTTGATTTGATCAAA 1440
 Db 1452 TTTTGTAAACATTAATATTTTGTACAGTTAAAGTAAATATTTGATTTGATCAAA 1511
 QY 1441 TAGCATATTAACCTTTGTAAGCAAGTATGCTTGTATGATGATTAAGTAAAGTAAATTAATTAAT 1500
 Db 1512 TAGCATATTAACCTTTGTAAGCAAGTATGCTTGTATGATGATTAAGTAAAGTAAATTAATTAAT 1571
 QY 1501 AATTTTCTTTTGAATTAACATTTTAAATPACCTTTGAAATATCCTTTGTGTCAGTG 1560
 Db 1572 AATTTTCTTTTGAATTAACATTTTAAATPACCTTTGAAATATCCTTTGTGTCAGTG 1631
 QY 1561 ATAAATGTGATGATCTTGGCTTTTGTACATGAGGTCACCTCTGAAGTATTTTGTG 1620
 Db 1632 ATAAATGTGATGATCTTGGCTTTTGTACATGAGGTCACCTCTGAAGTATTTTGTG 1691
 QY 1621 AGTAAAGGAAATCTTGACTACTTTATTTCTTAAAGGAATATCTTTATATACCTTCAAA 1680
 Db 1692 AGTAAAGGAAATCTTGACTACTTTATTTCTTAAAGGAATATCTTTATATACCTTCAAA 1751
 QY 1681 TTTTGAACCTTAACTTTAAAGTTTTTCTTCTGTAATTTGTGTAACGGGTGATTTATTATTA 1740
 Db 1752 TTTTGAACCTTAACTTTAAAGTTTTTCTTCTGTAATTTGTGTAACGGGTGATTTATTATTA 1811
 QY 1741 CTCTAGATAGCAGGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTGATGATCTAT 1800
 Db 1812 CTCTAGATAGCAGGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTGATGATCTAT 1871

QY 1801 TAAATTTTAAAGTGTGTATTTCTTTTCAATGGGTGATGTACGGGTGATAACAGACATTC 1860
 Db 1872 TAAATTTTAAAGTGTGTATTTCTTTTCAATGGGTGATGTACGGGTGATAACAGACATTC 1931
 QY 1861 ATGGAAGGCACTGCAAGTTTGTGTCATGTCAGAGTTTGTATTAATAAACACACATACACT 1920
 Db 1932 ATGGAAGGCACTGCAAGTTTGTGTCATGTCAGAGTTTGTATTAATAAACACACATACACT 1991
 QY 1921 TTAATTAAGATTAATAATCTAACTGGAAGTACGCTTTGGAAGTACGATTTTCAAGTATG 1980
 Db 1992 TTAATTAAGATTAATAATCTAACTGGAAGTACGCTTTGGAAGTACGATTTTCAAGTATG 2051
 QY 1981 TTTGTGAGTCAAGATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2040
 Db 2052 TTTGTGAGTCAAGATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2111
 QY 2041 AAGTCTTTAGGAGTCTTAACATTTGGAAGTACGCTTTGGAAGTACGATTTTCAAGTATG 2100
 Db 2112 AAGTCTTTAGGAGTCTTAACATTTGGAAGTACGCTTTGGAAGTACGATTTTCAAGTATG 2171
 QY 2101 ACCTATAAGAAATTAAGTTTATTAATTAAGTCAATTTATGCTGTGATTAATTTTACGGGAG 2160
 Db 2172 ACCTATAAGAAATTAAGTTTATTAATTAAGTCAATTTATGCTGTGATTAATTTTACGGGAG 2231
 QY 2161 AAAGAGGATTTGATTTGGAAGTACGCTTTGGAAGTACGCTTTGGAAGTACGATTTTCCAGAAATTT 2220
 Db 2232 AAAGAGGATTTGATTTGGAAGTACGCTTTGGAAGTACGCTTTGGAAGTACGATTTTCCAGAAATTT 2291
 QY 2221 AATGATTTGGTTTACATAAATCTTTTGAATTTGACTTTCAGAAA 2256
 Db 2292 AATGATTTGGTTTACATAAATCTTTTGAATTTGACTTTCAGAAA 2327

RESULT 2

US-10-111-076A-1
 ; Sequence 1, Application US/10111076A
 ; GENERAL INFORMATION:

; APPLICANT: XIE, YI
 ; TITLE OF INVENTION: A NEW POLYPEPTIDE - NEW CELL CYCLE-REGULATING PROTEIN 53 AND A
 ; FILE REFERENCE: 2310/51152
 ; CURRENT APPLICATION NUMBER: US/10/111,076A
 ; CURRENT FILING DATE: 2002-04-19
 ; PRIOR APPLICATION NUMBER: PCT/CN/00328
 ; PRIOR FILING DATE: 2000-10-16
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1

; LENGTH: 2497
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens

; FEATURES:
 ; NAME/KEY: CDS
 ; LOCATION: (111)..(1556)
 ; OTHER INFORMATION:

US-10-111-076A-1

Query Match 98.2%; Score 2215.8; DB 48; Length 2497;
 Best Local Similarity 99.7%; Pred. No. 0;

Matches 2220; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 14 GTCAATATTTAGAACGAGGTCCTTGAATGACGAGATTTATCGGACCGGAGATACGTTG 73
 Db 271 GTCAATATTTAGAACGAGGTCCTTGAATGACGAGATTTATCGGACCGGAGATACGTTG 330
 QY 74 ACGAATACAGGAATGACTACTCTGAAGGATATGTTCTTAGACATTTATCACAGACATTTG 133
 Db 331 ACGAATACAGGAATGACTACTCTGAAGGATATGTTCTTAGACATTTATCACAGGCAATG 390
 QY 134 AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCACTCGGACGAGGAGCACTCTTA 193

Db 391 AAAGGGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGCGAGCGAGAGCGAGTCCTA 450
Qy 194 AAAGGAAGCGCAATAGACACATGTTCAAGTCATCACTGTTCAAGTCACTGTTCAAGGAGCCACCGAAGGA 253
Db 451 AAAGGAAGCGCAATAGACACATGTTCAAGTCATCACTGTTCAAGTCACTGTTCAAGGAGCCACCGAAGGA 510
Qy 254 AAAGATCCAGGAGTATAGAGATAGATGATGAGAGGAGTCACTGATCTGTCAAAATGAGAGACG 313
Db 511 AAAGATCCAGGAGTATAGAGATGATGAGAGGAGTCACTGATCTGTCAAAATGAGAGACG 570
Qy 314 TTCTAAGAGCAAGATATGAATCGTGGACACATTTGGGTGAAGGAGCCCTTGGCAAAAGTG 373
Db 571 TTCTAAGAGCAAGATATGAATCGTGGACACATTTGGGTGAAGGAGCCCTTGGCAAAAGTG 630
Qy 374 TAGAGTGCAATGATCATCGGCATGAGTGCATGTCAGTACGAGTGAATATCGTAAATAATG 433
Db 631 TAGAGTGCAATGATCATCGGCATGAGTGCATGTCAGTACGAGTGAATATCGTAAATAATG 690
Qy 434 TAGGCGGTTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCACTTAAATAGTA 493
Db 691 TAGGCGGTTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCACTTAAATAGTA 750
Qy 494 CTGATCCCAATAGTCTTCGATGTCGATGTCGAGATGCTAGATGTTTTCATCATCATGTC 553
Db 751 CTGATCCCAATAGTCTTCGATGTCGATGTCGAGATGCTAGATGTTTTCATCATCATGTC 810
Qy 554 ATGTTTGTATGTTGTTCAACTACTGCGACCTTAGTACTACGATTCATTAAGAAAAACA 613
Db 811 ATGTTTGTATGTTGTTGAACTACTGCGACCTTAGTACTACGATTCATTAAGAAAAACA 870
Qy 614 GCTTTCGCCATTCCTCAATGACCAATCAGGAGATGCGGTATCAGATCTGCCAGTCAA 673
Db 871 GCTTTCGCCATTCCTCAATGACCAATCAGGAGATGCGGTATCAGATCTGCCAGTCAA 930
Qy 674 TABATTTTTTACATCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 733
Db 931 TABATTTTTTACATCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 990
Qy 734 TTGTGAAGTCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 793
Db 991 TTGTGAAGTCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1050
Qy 794 TGAATAACACAGATATCAAAAGTGTGTGATTTGGAAGTGCACGATGATGATGATGATGATGAT 853
Db 1051 TGAATAACACAGATATCAAAAGTGTGTGATTTGGAAGTGCACGATGATGATGATGATGATGAT 1110
Qy 854 ACAGTACTTTGGTGTCTACCGGCACTACAGAGTCCCGAGGTCAATTTGGCTTTAGGTT 913
Db 1111 ACAGTACTTTGGTGTCTACCGGCACTACAGAGTCCCGAGGTCAATTTGGCTTTAGGTT 1170
Qy 914 GGTCTCAGCCTTTGATGTTTGGAGCATAGGTTGCAATCTTATTGAAATATTACCTTTGGTT 973
Db 1171 GGTCTCAGCCTTTGATGTTTGGAGCATAGGTTGCAATCTTATTGAAATATTACCTTTGGTT 1230
Qy 974 TCACAGTCTTTTCAGACTCATGATAGTAAAGAGCACCTTGGCAATGATGGAACGAATATTAG 1033
Db 1231 TCACAGTCTTTTCAGACTCATGATAGTAAAGAGCACCTTGGCAATGATGGAACGAATATTAG 1290
Qy 1034 GACCCATACACACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1093
Db 1291 GACCCATACACACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1350
Qy 1094 AGCTAGATTGGATGCAACACAGTCTCTGCTGGTGTAGATATGTTAGGAGAGCGCTGCAACCGT 1153
Db 1351 AGCTAGATTGGATGCAACACAGTCTCTGCTGGTGTAGATATGTTAGGAGAGCGCTGCAACCGT 1410
Qy 1154 TGAAGGAATTTATGCTTTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1213
Db 1411 TGAAGGAATTTATGCTTTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1470
Qy 1214 GAATGTTAGATATGATCCCACTCAAGAAATTTACCTTGGATGCAAGCAATGAGCATCTCTT 1273
Db 1471 GAATGTTAGATATGATCCCACTCAAGAAATTTACCTTGGATGCAAGCAATGAGCATCTCTT 1530

RESULT 3
US-09-979-167-55
; Sequence 55, Application US/09979167

; GENERAL INFORMATION:
 ; APPLICANT: PLOWMAN, GREGORY D.
 ; APPLICANT: MARTINEZ, RICARDO
 ; APPLICANT: WHYTE, DAVID
 ; APPLICANT: SUDERSANAM, SUCHA
 ; TITLE OF INVENTION: PROTEIN KINASES
 ; FILE REFERENCE: 038602/1273
 ; CURRENT APPLICATION NUMBER: US/09/979,167
 ; CURRENT FILING DATE: 2001-11-20
 ; NUMBER OF SEQ ID NOS: 269
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 55
 ; LENGTH: 2488
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: modified_base
 ; LOCATION: (562)
 ; OTHER INFORMATION: a, t, c, g, other or unknown
 US-09-979-167-55

Query Match 90.8%; Score 2048; DB 44; Length 2488;
 Best Local Similarity 95.9%; Pred. No. 0;
 Matches 2148; Conservative 0; Mismatches 1; Indels 90; Gaps 1;

QY	14	GTCAATATTAGAACCAAGGTCCTTGAATGACGAGATTTATCGGACCGGAGATACGTTG	73
DB	338	GTCAATATTAGAACCAAGGTCCTTGAATGACGAGATTTATCGGACCGGAGATACGTTG	397
QY	74	ACGAATACAGGAATGACTACTGCTGAAGGATATGTTCTTAGACATTTATCACAGACATTTG	133
DB	398	ACGAATACAGGAATGACTACTGCTGAAGGATATGTTCTTAGACATTTATCACAGACATTTG	457
QY	134	AAAGCGGGTATCGAATCCACTCGCTCAATCTTCACTCGCAGCAGGAGGAGCAGTCCTA	193
DB	458	AAAGCGGGTATCGAATCCACTCGCTCAATCTTCACTCGCAGCAGGAGGAGCAGTCCTA	517
QY	194	AAAGGAAGCGCAATAGACACTGTTCAAGTCATCAGTCAGTTCGAAAGGCCACCGAAGGA	253
DB	518	AAAGGAAGCGCAATAGACACTGTTCAAGTCATCAGTCAGTTCGAAAGGCCACCGAAGGA	561
QY	254	AAAGATCCAGGAGTATAGGAGTATGAGGAGGTCACCTGATCTGTCAAAGTGAGAGC	313
DB	562	-----	561
QY	314	TTCTAAGACAGATATGAAATCGTGGACACTTTGSGTGAAGGAGCCTTTGSCAAAGTTG	373
DB	562	-----NATGAAATCGTGGACACTTTGSGTGAAGGAGCCTTTGSCAAAGTTG	607
QY	374	TAGAGTGCATTCATGCGCATGGATGGCATGTCATGTAGCAGTGAATAATCGTAAATAATG	433
DB	608	TAGAGTGCATTCATGCGCATGGATGGCATGTCATGTAGCAGTGAATAATCGTAAATAATG	667
QY	434	TAGGCGGTTACCGTGAAGCAGCTGTTTCAAGAAATCAAGATATAGAGCACTTAAATAGTA	493
DB	668	TAGGCGGTTACCGTGAAGCAGCTGTTTCAAGAAATCAAGATATAGAGCACTTAAATAGTA	727
QY	494	CTGATCCCAATAGTCTCTCCGATGTCAGATGCTAGAAATGTTTGCATCATCATGGTC	553
DB	728	CTGATCCCAATAGTCTCTCCGATGTCAGATGCTAGAAATGTTTGCATCATCATGGTC	787
QY	554	ATGTTTGTATTGTTTGAATCTACTGGGACTTTAGTACTTTACCATTTCAATTAAGAAAAACA	613
DB	788	ATGTTTGTATTGTTTGAATCTACTGGGACTTTAGTACTTTACCATTTCAATTAAGAAAAACA	847
QY	614	GCTTTCGCAATTCAAATGTACCAATCAGCAGATGGCGTATCAGATCTGCCAGTCAA	673
DB	848	GCTTTCGCAATTCAAATGTACCAATCAGCAGATGGCGTATCAGATCTGCCAGTCAA	907
QY	674	TAAATTTTACATCATATAATTAATTAACCATACAGATCTGAAGCTGAAATATTTTGT	733
DB	908	TAAATTTTACATCATATAATTAATTAACCATACAGATCTGAAGCTGAAATATTTTGT	967

QY	734	TTGTGAAGTCTGACTATATGTAGTCAAAATATAATTTTAAATGAAACGTGATGAACGACAC	793
DB	968	TTGTGAAGTCTGACTATATGTAGTCAAAATATAATTTTAAATGAAACGTGATGAACGACAC	1027
QY	794	TGAAAAACACAGATATCAAAAGTTGTGACTTTTGGAAAGTGAACGATGATGATGAACATC	853
DB	1028	TGAAAAACACAGATATCAAAAGTTGTGACTTTTGGAAAGTGAACGATGATGATGAACATC	1087
QY	854	ACAGTACTTTTGGTGTCTACCCGGCACTACAGAGCTCCCGAGGTCATTTTGGGCTTTAGGTT	913
DB	1088	ACAGTACTTTTGGTGTCTACCCGGCACTACAGAGCTCCCGAGGTCATTTTGGGCTTTAGGTT	1147
QY	914	GGTCTCAGCCTTTGATGTTTGGAGCATAGTTGCAATCTTTATTTGAATATTAACCTTGGTT	973
DB	1148	GGTCTCAGCCTTTGATGTTTGGAGCATAGTTGCAATCTTTATTTGAATATTAACCTTGGTT	1207
QY	974	TCACAGTCTTTTCAGACTCATATAGTAAAGAGCACCTGCAATGATGAACCAATATTAG	1033
DB	1208	TCACAGTCTTTTCAGACTCATATAGTAAAGAGCACCTGCAATGATGAACCAATATTAG	1267
QY	1034	GACCCATACCAACACACATGATTCAGAAAAACAAGAAAAACGCAAGTATTTTCCACATAACC	1093
DB	1268	GACCCATACCAACACACATGATTCAGAAAAACAAGAAAAACGCAAGTATTTTCCACATAACC	1327
QY	1094	AGCTAGATGGGATGAACACACAGTCTCTGCTGTGTAGATATGTTAGGAGACGCTGCCAACCGT	1153
DB	1328	AGCTAGATGGGATGAACACACAGTCTCTGCTGTGTAGATATGTTAGGAGACGCTGCCAACCGT	1387
QY	1154	TGAAGGAATTTATGCTTTGTCATGATGAAGACATGAGAACTGTTTTCACCTGGTTCGAA	1213
DB	1388	TGAAGGAATTTATGCTTTGTCATGATGAAGACATGAGAACTGTTTTCACCTGGTTCGAA	1447
QY	1214	GAACTTTAGATATATGATCCAACTCAAAGAAATTTACCTTGGATGAAGCAATTCAGCATCCTT	1273
DB	1448	GAACTTTAGATATATGATCCAACTCAAAGAAATTTACCTTGGATGAAGCAATTCAGCATCCTT	1507
QY	1274	TCTTTGACTTATTTAAAAAGAAATGAAATGGAAATCAGTGGTCTTACTATATATCTCTCT	1333
DB	1508	TCTTTGACTTATTTAAAAAGAAATGAAATGGAAATCAGTGGTCTTACTATATATCTCTCT	1567
QY	1334	AGAAGAGTACTTTAAGACTGTGTCAGTCACTCACTAAACATTTCTAATATTTTGTAAACAT	1393
DB	1568	AGAAGAGTACTTTAAGACTGTGTCAGTCACTCACTAAACATTTCTAATATTTTGTAAACAT	1627
QY	1394	AAATTTATTTGTACAGTTAAGTGAATATTTGTATGTTTGTATCAATAGCAATTAATAC	1453
DB	1628	AAATTTATTTGTACAGTTAAGTGAATATTTGTATGTTTGTATCAATAGCAATTAATAC	1687
QY	1454	TTGTTAAGCAAGTATGGTCTTGAATATGCAATTTAGAAAAATTTAAATTTCTTTT	1513
DB	1688	TTGTTAAGCAAGTATGGTCTTGAATATGCAATTTAGAAAAATTTAAATTTCTTTT	1747
QY	1514	GAAATACATTTTAAATACCTTTGAAATATCTTTTGTGTCAGTGAATATGATG	1573
DB	1748	GAAATACATTTTAAATACCTTTGAAATATCTTTTGTGTCAGTGAATATGATG	1807
QY	1574	ATCTTGCTTTTGTACATGAGGTCACCTCTGAAAGTGAATTTTGTAGTAAAGAAAT	1633
DB	1808	ATCTTGCTTTTGTACATGAGGTCACCTCTGAAAGTGAATTTTGTAGTAAAGAAAT	1867
QY	1634	CTTGACTATTTTATATCTTAAAGGAATATCTTTTATATATCTTCAAAATTTAGAACTAAC	1693
DB	1868	CTTGACTATTTTATATCTTAAAGGAATATCTTTTATATATCTTCAAAATTTAGAACTAAC	1927
QY	1694	TTTAAAGTCTTTCTCTGTAAATTTGTTGAAACGGGTATTTATTTAACTCTAGATAAGCA	1753
DB	1928	TTTAAAGTCTTTCTCTGTAAATTTGTTGAAACGGGTATTTATTTAACTCTAGATAAGCA	1987
QY	1754	GGTACTAGAAACCAAACTCAGAAATGTTTACTCTGTAGAAATTTCTATTTAAATTTTAAAGT	1813
DB	1988	GGTACTAGAAACCAAACTCAGAAATGTTTACTCTGTAGAAATTTCTATTTAAATTTTAAAGT	2047
QY	1814	TTGTATTTCTTTTCAATTTGGGTGATGTCAGGGGTGATAACCAAGACATTTTCATGGAAAGGCATG	1873


```

Db 1463 TCTTTGACTTATTAATAAAGAAATCAATGGAATCAGTGGTCTTACTATATCTTCTCT 1522
Qy 1334 AGAAGAGATTAATTAAGACTGTGTGAGTCAATCAATCAATCTTAATATTTTGTAAACAT 1393
Db 1523 AGAAGAGATTAATTAAGACTGTGTGAGTCAATCAATCAATCTTAATATTTTGTAAACAT 1582
Qy 1394 AAATTAATTTGTACAGTTAAGTGTAAATATTTGTATGTTTGTATCAATAGCATAATTAAC 1453
Db 1583 AAATTAATTTGTACAGTTAAGTGTAAATATTTGTATGTTTGTATCAATAGCATAATTAAC 1642
Qy 1454 TTGTTAAGCAAGTATGTTCTTGATATGATGATGATGATGATGATGATGATGATGATGATG 1513
Db 1643 TTGTTAAGCAAGTATGTTCTTGATATGATGATGATGATGATGATGATGATGATGATGAT 1702
Qy 1514 GAAATTAACATTTTAAATACCTTTGAAATATCTTTGATGATGATGATGATGATGATGATG 1573
Db 1703 GAAATTAACATTTTAAATACCTTTGAAATATCTTTGATGATGATGATGATGATGATGATG 1762
Qy 1574 ATCTTGCCCTTTTGTACATGAGGTCACTCTGAGTGAATTTTGTGATGATGATGATGATG 1633
Db 1763 ATCTTGCCCTTTTGTACATGAGGTCACTCTGAGTGAATTTTGTGATGATGATGATGATG 1822
Qy 1634 CTGTGACTACTTATATCTTAAAGGAATATCTTTATATATCTTCAATTTTGAACCTTAAC 1693
Db 1823 CTGTGACTACTTATATCTTAAAGGAATATCTTTATATATCTTCAATTTTGAACCTTAAC 1882
Qy 1694 TTTAAAGTTTTTCTCTGTAATTTGTAAGGATGATTTATTAATTAATCTAGATAAGCA 1753
Db 1883 TTTAAAGTTTTTCTCTGTAATTTGTAAGGATGATTTATTAATTAATCTAGATAAGCA 1942
Qy 1754 GGTACTAGAAACCAAACTCAGAAATGTTTACTGTTAGAAATCTATTAATTTTAAAGTG 1813
Db 1943 GGTACTAGAAACCAAACTCAGAAATGTTTACTGTTAGAAATCTATTAATTTTAAAGTG 2002
Qy 1814 TTGTATCTCTTTTCAATTCAGTGTGATGTCAGGTTGATTAACAGACATCTATGTAAGGCA 1873
Db 2002 TTGTATCTCTTTTCAATTCAGTGTGATGTCAGGTTGATTAACAGACATCTATGTAAGGCA 2062
Qy 1874 CAGTTTGTCCATGTTGACAGTTTGTTTAATAAACCACATACACACATTTTAAAGATTA 1933
Db 2063 CAGTTTGTCCATGTTGACAGTTTGTTTAATAAACCACATACACACATTTTAAAGATTA 2122
Qy 1934 AAATCTAACTGGAAGTCACTGTGAAATGGAATGGAATTTCCAGATATGTTGGTGAAGTCA 1993
Db 2123 AAATCTAACTGGAAGTCACTGTGAAATGGAATTTCCAGATATGTTGGTGAAGTCA 2182
Qy 1994 AGATATAAAATAGAAATTTCTGATGAGAGTTTCAAGTTTAAATACCAAGTCTTTAGGAG 2053
Db 2183 AGATATAAAATAGAAATTTCTGATGAGAGTTTCAAGTTTAAATACCAAGTCTTTAGGAG 2242
Qy 2054 TCTTAAACATGGCCAGCATCTGTTTATCAATGACATAAATACGTAACCTTATAAGAAAT 2113
Db 2243 TCTTAAACATGGCCAGCATCTGTTTATCAATGACATAAATACGTAACCTTATAAGAAAT 2302
Qy 2114 AAGTTTATTAATAGGCAATTTATGTCGTGATATTTCTTACGGGAGAAAGAGGATTTGA 2173
Db 2303 AAGTTTATTAATAGGCAATTTATGTCGTGATATTTCTTACGGGAGAAAGAGGATTTGA 2362
Qy 2174 TTGGAAGCAGTTTGGGAGAAAGTCTGCTGAAATTTCCAGATTTTAAATTTGATTTGTTA 2233
Db 2363 TTGGAAGCAGTTTGGGAGAAAGTCTGCTGAAATTTCCAGATTTTAAATTTGATTTGTTA 2422
Qy 2234 CATAAACCTTTTGTACTTCA 2252
Db 2423 CATAAACCTTTTGTACTTCA 2441

```

RESULT 5
 US-60-278-3721
 ; Sequence 3721, Application US/60278258
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, MacDonald
 ; APPLICANT: Lal, Preeti

; APPLICANT: Diep, Dinh
 ; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
 ; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
 ; TITLE OF INVENTION: Polymorphisms Identified Thereby
 ; FILE REFERENCE: GX-0010-1 P
 ; CURRENT APPLICATION NUMBER: US/60/278,258
 ; CURRENT FILING DATE: 2001-03-23
 ; NUMBER OF SEQ ID NOS: 17730
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 3721
 ; LENGTH: 2487
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No: 253847.3
 US-60-278-3721

Query Match	90.7%	Score	2047	DB	97	Length	2487
Best Local Similarity	95.9%	Pred. No.	0				
Matches	2148	Conservative	0	Mismatches	0	Indels	91
Gaps	1						

Qy	14	GTCAATATTTTAAAGCAAGTCCCTTGAATGAGCGAGATTTATCGGACCGGAGATACGTTG	73
Db	338	GTCAATATTTTAAAGCAAGTCCCTTGAATGAGCGAGATTTATCGGACCGGAGATACGTTG	397
Qy	74	ACGAATACAGGAATGACTTACTGTGAAGGATATGTTCTTAGACATTATCACAGAGACATTG	133
Db	398	ACGAATACAGGAATGACTTACTGTGAAGGATATGTTCTTAGACATTATCACAGAGACATTG	457
Qy	134	AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGAGCAGGAGAGAGCAGTCCTA	193
Db	458	AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGAGCAGGAGAGAGCAGTCCTA	517
Qy	194	AAAGGAGCGCAATAGACACATGTTCAAGTCACTAGTCACTGTCGAAGAGCCACCGAAGGA	253
Db	518	AAAGGAGCGCAATAGACACATGTTCAAGTCACTAGTCACTGTCGAAGAGCCACCGAAGGA	561
Qy	254	AAAGTCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAGTGGAGAGC	313
Db	562	-----	561
Qy	314	TTCTAAGAGCAAGATATGAATTCGTGGACACTTTGGGTGAAGAGGCCCTTTGGCAAGTTG	373
Db	562	-----ATGAATTCGTGGACACTTTGGGTGAAGAGGCCCTTTGGCAAGTTG	606
Qy	374	TAGAGTGCATTTGATCATGGCATGGATGCGATGTCATGTCAGTGAATTCGTAAATAATG	433
Db	607	TAGAGTGCATTTGATCATGGCATGGATGCGATGTCATGTCAGTGAATTCGTAAATAATG	666
Qy	434	TAGGCCGTTACCGTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTTAGAGCACCTTAAATAGTA	493
Db	667	TAGGCCGTTACCGTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTTAGAGCACCTTAAATAGTA	726
Qy	494	CTGATCCCAATAGTGTCTTCGATGTCGAGATGTCAGATGTCAGATGTCATCATGTC	553
Db	727	CTGATCCCAATAGTGTCTTCGATGTCGAGATGTCAGATGTCAGATGTCATCATGTC	786
Qy	554	ATGTTTGTATTTGTTGAACTACTGGGACCTAGTACTTACGATTTTCAATTAAGAAACA	613
Db	787	ATGTTTGTATTTGTTGAACTACTGGGACCTAGTACTTACGATTTTCAATTAAGAAACA	846
Qy	614	GCTTTCTGCCATTTCAAATTCACCAATCCAGATGCGCGTATCAGATCTGCCAGTCAA	673
Db	847	GCTTTCTGCCATTTCAAATTCACCAATCCAGATGCGCGTATCAGATCTGCCAGTCAA	906
Qy	674	TAAATTTTTCATCAATAATAAATTAACCCATAAGATCTGAAAGCCTGAAATATTTTGT	733
Db	907	TAAATTTTTCATCAATAATAAATTAACCCATAAGATCTGAAAGCCTGAAATATTTTGT	966
Qy	734	TTGTGAAGTCTGACTATGTTAGTCAATATAATTTCTAAATTAAGAGCGATGACGACAC	793
Db	967	TTGTGAAGTCTGACTATGTTAGTCAATATAATTTCTAAATTAAGAGCGATGACGACAC	1026


```
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: NEUTEMT01
; CLONE: 33963
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-392-978-9

Query Match          90.7%; Score 2046.6; DB 21; Length 2446;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 2149; Conservative 0; Mismatches 4; Indels 90; Gaps 1;

QY 14 GTCATATTAGAACCAAGGTCCTTGAATGACGAGATATTATCGGACCGGAGATACGTTG 73
DB 288 GTCATATTATTAGAACCAAGGTCCTTGAATGACGAGATATTATCGGACCGGAGATACGTTG 347
QY 74 ACGAATACAGGAATGACTACTCTGTAAGGATATGTTCTTAGACATTATCACAGAGACATTG 133
DB 348 ACGAATACAGGAATGACTACTCTGTAAGGATATGTTCTTAGACATTATCACAGAGACATTG 407
QY 134 AAAGCGGGTATCGAATCCACTGACGTAAATCTTCAGTCCGACGAGGAGAGACAGTCCCTA 193
DB 408 AAAGCGGGTATCGAATCCACTGACGTAAATCTTCAGTCCGACGAGGAGAGACAGTCCCTA 467
QY 194 AAAGGAAGCGGCAATAGACACTGTTCAAGTCATCATGCTGAGAGCCACCGAAGGA 253
DB 468 AAAGGAAGCGGCAATAGACACTGTTCAAGTCATCATGCTGAGAGCCACCGAAGGA 515
QY 254 AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACTGATCTGTCMAAGTGGAGAG 313
DB 516 ----- 515
QY 314 TTCTAAGACCAAGATATGAATCGTGGACACCTTTGGGTGAAGGAGCCTTTGGCAAGTTG 373
DB 516 -----AAATCGGTGGACACTTTGGGTGAGGAGCCTTTGGCAAGTTG 557
QY 374 TAGAGTGCAATTGATCATGGCATGGAATGCGATGTCATGATGAGGAGTGAATTCGTAAGG 433
DB 558 TAGAGTGCAATTGATCATGGCATGGAATGCGATGTCATGATGAGGAGTGAATTCGTAAGG 617
QY 434 TAGGCGGTACCGTGAAGCAGTCTGTTGAGAAATCCAAATCCAAATCCAAATCCAAATAGTA 493
DB 618 TAGGCGGTACCGTGAAGCAGTCTGTTGAGAAATCCAAATCCAAATCCAAATCCAAATAGTA 677
QY 494 CTGATCCCAATAGTGTCTCCGATGCTGTCAGATGCTGAGATGTTGATCATCATGTC 553
DB 678 CTGATCCCAATAGTGTCTCCGATGCTGTCAGATGCTGAGATGTTGATCATCATGTC 737
QY 554 ATGTTTGTATTGTTGAACTACTGGGACTTAGTACTTACGATTTCAATTAAGAAAAACA 613
DB 738 ATGTTTGTATTGTTGAACTACTGGGACTTAGTACTTACGATTTCAATTAAGAAAAACA 797
QY 614 GCTTTCTGCCATTTCAAATTTGACCATCATCAGGACATGCGGTATCAGATCTGCCAGTCAA 673
DB 798 GCTTTCTGCCATTTCAAATTTGACCATCATCAGGACATGCGGTATCAGATCTGCCAGTCAA 857
QY 674 TAAATTTTTTATCATATAATTAATTAACCATPACAGATCTGAAAGCTGAAAATATTTTGT 733
DB 858 TAAATTTTTTATCATATAATTAATTAACCATPACAGATCTGAAAGCTGAAAATATTTTGT 917
QY 734 TTGTGAAGTCTGACTATGATGATCAATATAATTTCTTAAATGAAACGTCATGATGAAACATC 793
DB 918 TTGTGAAGTCTGACTATGATGATCAATATAATTTCTTAAATGAAACGTCATGATGAAACATC 977
QY 794 TGAATAACACAGATATCAAAGTTGTTGACTTTGGAAAGTGCACGTCATGATGAAACATC 853
DB 978 TGAATAACACAGATATCAAAGTTGTTGACTTTGGAAAGTGCACGTCATGATGAAACATC 1037
QY 854 ACAGTACTTTGGTGTCTACCGGCACTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGGTT 913
DB 1038 ACAGTACTTTGGTGTCTACCGGCACTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGGTT 1097
QY 914 GGTCTCAGCCTTGTCATGTTTGGAGCATAGGTTGCAATCTTATTGAATATTACCTTGGTT 973
```

```
DB 1098 GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCAATCTTATTGAATATTACCTTGGTT 1157
QY 974 TCACAGTCTTTTCAGACTCATGATAGTAAGAGCAGCTTGGCAATGATGGAACGAATATTAG 1033
DB 1158 TCACAGTCTTTTCAGACTCATGATAGTAAGAGCAGCTTGGCAATGATGGAACGAATATTAG 1217
QY 1034 GACCCATACCAACAACATGATTTCAAGAAAAAACAAGAAACGCAAGTATTTTTACCATAACC 1093
DB 1218 GACCCATACCAACAACATGATTTCAAGAAAAAACAAGAAACGCAAGTATTTTTACCATAACC 1277
QY 1094 AGCTAGATTGGGATGAACAACAAGTCTGCTGCTAGATATGTTAGGAGAGCGCTGCAACCGT 1153
DB 1278 AGCTAGATTGGGATGAACAACAAGTCTGCTGCTAGATATGTTAGGAGAGCGCTGCAACCGT 1337
QY 1154 TGAAGGAATTTATGCTTTGTGTCATGATGAAGAAACATGAGAAACGTTTGCACCTGGTTCCGAA 1213
DB 1338 TGAAGGAATTTATGCTTTGTGTCATGATGAAGAAACATGAGAAACGTTTGCACCTGGTTCCGAA 1397
QY 1214 GAATGTTAGATATGATCCAACTCAAAGAAATTAACCTTGGATGAAGCATTCGACGATCCCT 1273
DB 1398 GAATGTTAGATATGATCCAACTCAAAGAAATTAACCTTGGATGAAGCATTCGACGATCCCT 1457
QY 1274 TCTTTGACTTATTAAGAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATCTTCTCT 1333
DB 1458 TCTTTGACTTATTAAGAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATCTTCTCT 1517
QY 1334 AGAAGAGATTAATTAAGAGCTGTGTCAGTCAACTAAACATTTCTAAATTTTGTAAACAT 1393
DB 1518 AGAAGAGATTAATTAAGAGCTGTGTCAGTCAACTAAACATTTCTAAATTTTGTAAACAT 1577
QY 1394 AAATTAATTTGTACAGTTAAGTGAATTAATTTGTATGTTTGTATCAATAGCATATAATTAAC 1453
DB 1578 AAATTAATTTGTACAGTTAAGTGAATTAATTTGTATGTTTGTATCAATAGCATATAATTAAC 1637
QY 1454 TTGTTAAGCAATGATGCTTGTGATTAATGCAATAGAAAAATTAATAATTTTCTTCTTTT 1513
DB 1638 TTGTTAAGCAATGATGCTTGTGATTAATGCAATAGAAAAATTAATAATTTTCTTCTTTT 1697
QY 1514 GAAATTAACATTTTAAATACCTTTGAAATATCTTTGTGTCAGTGAATATGATGTTG 1573
DB 1698 GAAATTAACATTTTAAATACCTTTGAAATATCTTTGTGTCAGTGAATATGATGTTG 1757
QY 1574 ATCTTGCTTTTGTACATGAGAGTCACTCTGAGTGAATTTTTTTTGTAGTAAAGGAAT 1633
DB 1758 ATCTTGCTTTTGTACATGAGAGTCACTCTGAGTGAATTTTTTTTGTAGTAAAGGAAT 1817
QY 1634 CTGACTACTTTTATATTCTTAAAGGAATATTCTTTATATATCTTCAAAATTTAGAACTTAAC 1693
DB 1818 CTGACTACTTTTATATTCTTAAAGGAATATTCTTTATATATCTTCAAAATTTAGAACTTAAC 1877
QY 1694 TTTTAAAGTTTTTCTCTGTAATTTGTCGAACGGGTGATTATTATTAACTCTAGATAAGCA 1753
DB 1878 TTTTAAAGTTTTTCTCTGTAATTTGTCGAACGGGTGATTATTATTAACTCTAGATAAGCA 1937
QY 1754 GGTACTAGAAACCAAACTCAGAAAAATCTTACTCTTGTAGAAATCTTATTAATTTTAAAGTG 1813
DB 1938 GGTACTAGAAACCAAACTCAGAAAAATCTTACTCTTGTAGAAATCTTATTAATTTTAAAGTG 1997
QY 1814 TTGTATTTCTTTTCAATTTGGGTGATGTCAGGTGATTAACAGACATTCATGGAAGGCGATG 1873
DB 1998 TTGTATTTCTTTTCAATTTGGGTGATGTCAGGTGATTAACAGACATTCATGGAAGGCGATG 2057
QY 1874 CAGTTTGTCCATTTGTGACAGTTTGTGTTTAAATAAACCAACATACACACTTTTATTAGATTA 1933
DB 2058 CAGTTTGTCCATTTGTGACAGTTTGTGTTTAAATAAACCAACATACACACTTTTATTAGATTA 2117
QY 1934 AAATCTAACTGGAAGTCAAGCTTGGAAAAATGACATTTTCCAAAGTATGTTTGGTGAAGTCAAC 1993
DB 2118 AAATCTAACTGGAAGTCAAGCTTGGAAAAATGACATTTTCCAAAGTATGTTTGGTGAAGTCAAC 2177
QY 1994 AGATATAAAAAATAGAAATTTCTGATGAGAGGTTTTCAGTTTAAATACCAAGTCTTTAGGAG 2053
```

```

Db      2178 AGATATAAATAAGAAATCTGATGAGAGGTTTCAGTTTTTAAATACCAAGTCCTTAGGAG 2237
Qy      2054 TCTTAAACATGGCCAGCATCTGTTTATCAATAGACATAAATACGTAACCTTAAAGAAAT 2113
Db      2238 TCTTAAACATGGCCAGCATCTGTTTATCAATAGACATAAATACGTAACCTTAAAGAAAT 2297
Qy      2114 AAGTTTATTAATTAGGCAATTTATGTCTGTGATAATCTTACGGGAGAAAGAGGATTTGA 2173
Db      2298 AAGTTTATTAATTAGGCAATTTATGTCTGTGATAATCTTACGGGAGAAAGAGGATTTGA 2357
Qy      2174 TTGGAACAGTTTGGGAGAAAGTCTGCTGAAATTTCCAGAAATTTAAATGATTTGTTA 2233
Db      2358 TTGGAACAGTTTGGGAGAAAGTCTGCTGAAATTTCCAGAAATTTAAATGATTTGTTA 2417
Qy      2234 CATAACTTTTTTGACTTCCAGAAA 2256
Db      2418 CATAACTTTTTTGACTTCCAAAAA 2440

RESULT 7
US-10-425-114-26852
; Sequence 26852, Application US/10425114
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26852
; LENGTH: 4035
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4654-056-G4_FLI
US-10-425-114-26852

Query Match      84.6%; Score 1908.4; DB 54; Length 4035;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1909; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      327 ATATGAATCGTGGACACTTTGGGTGAAGAGGACCTTTGGCAAAAGTTGTAGAGTGCAATTGA 386
Db      2126 AGATGAATCGTGGACACTTTGGGTGAAGAGGACCTTTGGCAAAAGTTGTAGAGTGCAATTGA 2185
Qy      387 TCATGGCATGGATGGCATGGATGAGCAGTGAATATCGTAAATCGTAAATATGAGCCGTTACCG 446
Db      2186 TCATGGCATGGATGGCATGGATGAGCAGTGAATATCGTAAATCGTAAATATGAGCCGTTACCG 2245
Qy      447 TGAACAGCTCGTTCAGAAATCCAAAGTATTAGAGCACTTAAATAGTACTGATCCCAATAG 506
Db      2246 TGAACAGCTCGTTCAGAAATCCAAAGTATTAGAGCACTTAAATAGTACTGATCCCAATAG 2305
Qy      507 TGTCTTCGATGTGCCAGATGCTAGATGGTTTGATCATCATGGTGCATGGTTGTTGTTATGTT 566
Db      2306 TGTCTTCGATGTGCCAGATGCTAGATGGTTTGATCATCATGGTGCATGGTTGTTGTTATGTT 2365
Qy      567 GTTTGAACTACTGGCATTTAGTACTTACGATTTCAATTAAGAAACAGCTTTCTGCCATT 626
Db      2366 GTTTGAACTACTGGCATTTAGTACTTACGATTTCAATTAAGAAACAGCTTTCTGCCATT 2425
Qy      627 TCAAAATGACCAATCAGGAGATGGCGTATCAGATCTGCCAGTCAATAAATTTTTTACA 686
Db      2426 TCAAAATGACCAATCAGGAGATGGCGTATCAGATCTGCCAGTCAATAAATTTTTTACA 2485
Qy      687 TCATTAATTAATTAACCCATACAGATCTGAGGCTGAAATATTTTGTGTGAGAGTCTGA 746

```

```

Db      2486 TCATAATAAATAAATACCCATACAGATCTGAAGCCTGAAGAAATATTTTGTGTGAAGTCTGA 2545
Qy      747 CTATGTAGTCAAAATATAATTTCTAAATGAAACGATGAAGCGCACACTGTAAGAAACACAGA 806
Db      2546 CTATGTAGTCAAAATATAATTTCTAAATGAAACGATGAAGCGCACACTGTAAGAAACACAGA 2605
Qy      807 TATCAAAAGTTGTTGACTTTGGAGAGTGCAACGATATGATGATGAACATACACAGTACTTTGGT 866
Db      2606 TATCAAAAGTTGTTGACTTTGGAGAGTGCAACGATATGATGATGAACATACACAGTACTTTGGT 2665
Qy      867 GTCTACCCGGCAGTACAGAGCTCCGAGGTCATTTTGGCTTTAGGTTGGTCTCAGCCCTTG 926
Db      2666 GTCTACCCGGCAGTACAGAGCTCCGAGGTCATTTTGGCTTTAGGTTGGTCTCAGCCCTTG 2725
Qy      927 TGATGTTTGGAGCAGTAGTGTGCAATTTTATGAATATTTACCTTGGTTTTCACAGTCTTTTCA 986
Db      2726 TGATGTTTGGAGCAGTAGTGTGCAATTTTATGAATATTTACCTTGGTTTTCACAGTCTTTCA 2785
Qy      987 GACTCATGATAGTAAAGAGCACCTGGCAATGATGGAACGAATATTAGGACCCATACCA 1046
Db      2786 GACTCATGATAGTAAAGAGCACCTGGCAATGATGGAACGAATATTAGGACCCATACCA 2845
Qy      1047 ACACATGATTCAGAAACAGAAACCGNAGTATTTTACCACATACCGAGCTAGATGGGA 1106
Db      2846 ACACATGATTCAGAAACAGAAACCGNAGTATTTTACCACATACCGAGCTAGATGGGA 2905
Qy      1107 TGAACACAGTTCTGCTGCTAGATATGTTAGGAGACGCTGCAACCGTTGGAAGGAATTTAT 1166
Db      2906 TGAACACAGTTCTGCTGCTAGATATGTTAGGAGACGCTGCAACCGTTGGAAGGAATTTAT 2965
Qy      1167 GCTTTGTCATGATGAAGAACATGAGAACTGTTTGGCTGAGCTGTTGGAAGAAATGTTAGAATA 1226
Db      2966 GCTTTGTCATGATGAAGAACATGAGAACTGTTTGGCTGAGCTGTTGGAAGAAATGTTAGAATA 3025
Qy      1227 TGATCCAACTCAAGAAATACCTTGGATGAGAGCTGAGAGCATCTTTCTTTGACTTAT 1286
Db      3026 TGATCCAACTCAAGAAATACCTTGGATGAGAGCTGAGAGCATCTTTCTTTGACTTAT 3085
Qy      1287 AAAAAAGAAATGAAATGGGAATCAGTGTCTTACTATATATCTCTCTAGAGAGAGATTACT 1346
Db      3086 AAAAAAGAAATGAAATGGGAATCAGTGTCTTACTATATATCTCTCTAGAGAGAGATTACT 3145
Qy      1347 TAAGACTGTGTGCTCAACTAAACATCTCTAATATTTTGTGAAACATTAATAATTTTGTGA 1406
Db      3146 TAAGACTGTGTGCTCAACTAAACATCTCTAATATTTTGTGAAACATTAATAATTTTGTGA 3205
Qy      1407 CAGTTAAGTGTAAATATTTGATGTTTGTATCAATAGCATAAATTAACCTGTTTAAGCAAGT 1466
Db      3206 CAGTTAAGTGTAAATATTTGATGTTTGTATCAATAGCATAAATTAACCTGTTTAAGCAAGT 3265
Qy      1467 ATGCTCTTGATAATGCAATAGAAAAATTAATAATTTTCTTTTGTGAAATTTACCAAT 1526
Db      3266 ATGCTCTTGATAATGCAATAGAAAAATTAATAATTTTCTTTTGTGAAATTTACCAAT 3325
Qy      1527 TTAATACCTTTTGAATAATCTTTGTGTCAGTGATTAATGATGATGATCTTGCCTTTTG 1586
Db      3326 TTAATACCTTTTGAATAATCTTTGTGTCAGTGATTAATGATGATGATCTTGCCTTTTG 3385
Qy      1587 TACATGGAGGTCACCTCTGAAGTGAATTTTTCAGTAAAGGAAATCTTCACTACTTTTA 1646
Db      3386 TACATGGAGGTCACCTCTGAAGTGAATTTTTCAGTAAAGGAAATCTTCACTACTTTTA 3445
Qy      1647 TATTCTTAAAGGAATATTTCTTTATATATCTTCAAAATTTAGAACTTTAACTTTAAAGTTT 1706
Db      3446 TATTCTTAAAGGAATATTTCTTTATATATCTTCAAAATTTAGAACTTTAACTTTAAAGTTT 3505
Qy      1707 CTTCTGTAATTTGTAACCGGTGATTTATTAATCTCTAGATTAAGCAGGTACTAGAAACC 1766
Db      3506 CTTCTGTAATTTGTAACCGGTGATTTATTAATCTCTAGATTAAGCAGGTACTAGAAACC 3565
Qy      1767 AAAACTCAGAAAAATGTTTACTGTTAGAAATCTTATTAATTTTAAAGTTTGTATTTCTTTT 1826
Db      3566 AAAACTCAGAAAAATGTTTACTGTTAGAAATCTTATTAATTTTAAAGTTTGTATTTCTTTT 3625

```

QY 1827 CATTGGGTGATGTCAGGGTGATAAACAAGATTCATGGAAGGCATGCGATTGTGCCATT 1886
DB 3626 CATTGGGTGATGTCAGGGTGATAAACAAGATTCATGGAAGGCATGCGATTGTGCCATT 3685
QY 1887 GTGACAGTTTGTATTAATAAACACATACACACTTTTATTAAGATTAATACTAATCGGA 1946
DB 3686 GTGACAGTTTGTATTAATAAACACATACACACTTTTATTAAGATTAATACTAATCGGA 3745
QY 1947 AAGTCAGCTTGGAATAAGGACATTTCCAAAGTATGTTGGTGAAGTCAAGATTAATAAATA 2006
DB 3746 AAGTCAGCTTGGAATAAGGACATTTCCAAAGTATGTTGGTGAAGTCAAGATTAATAAATA 3805
QY 2007 GAAATCTGATGAGAGTTTCAGTTTATTAATACCAAGTCTTGAAGTCTTAACATTTGCG 2066
DB 3806 GAAATCTGATGAGAGTTTCAGTTTATTAATACCAAGTCTTGAAGTCTTAACATTTGCG 3865
QY 2067 CAGCATCTGTTTATCAATGACATTAATACGTTAAACCTTATAGATTAAGTTTATTAATT 2126
DB 3866 CAGCATCTGTTTATCAATGACATTAATACGTTAAACCTTATAGATTAAGTTTATTAATT 3925
QY 2127 AGGCAATTTATGTCGTGATTAATCTTACGGGAGAAAGAGATTTGATTGGAAGCAGTT 2186
DB 3926 AGGCAATTTATGTCGTGATTAATCTTACGGGAGAAAGAGATTTGATTGGAAGCAGTT 3985
QY 2187 TGGGAAGAAAGTGTCTGAAATTTCCAGAAATTTAAATGATTGTTACAT 2236
DB 3986 TGGGAAGAAAGTGTCTGAAATTTCCAGAAATTTAAATGATTGTTACAT 4035

RESULT 8
US-10-425-114A-26852
; Sequence 26852, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26852
; LENGTH: 4035
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4654-056-G4_FLI
US-10-425-114A-26852

Query Match 84.6%; Score 1908.4; DB 54; Length 4035;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1909; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 327 ATATGAATCTGGACACTTTGGGTGAAGGACCTTTGGCAAAAGTTGTAGAGTGCAATGA 386
DB 2126 AGATGAATCTGGACACTTTGGGTGAAGGACCTTTGGCAAAAGTTGTAGAGTGCAATGA 2185
QY 387 TCATGGCATGATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 446
DB 2186 TCATGGCATGATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2245
QY 447 TGAACAGCTCGTTGAGAAATCCAAAGTATTAGAGCACTTAATAGTACTGATCCCAATAG 506
DB 2246 TGAACAGCTCGTTGAGAAATCCAAAGTATTAGAGCACTTAATAGTACTGATCCCAATAG 2305
QY 507 TGTCTCCGATGTGTCAGATGCTGAGATGTTGATCATCATGATGATGATGATGATGATGATGAT 566
DB 2306 TGTCTCCGATGTGTCAGATGCTGAGATGTTGATCATCATGATGATGATGATGATGATGATGAT 2365

QY 567 GTTTGAACCTACTGGGACTTAGTACTTACGATTTCAATTAAGAAAAACAGCTTTCTGCCATT 626
DB 2366 GTTTGAACCTACTGGGACTTAGTACTTACGATTTCAATTAAGAAAAACAGCTTTCTGCCATT 2425
QY 627 TCATAATGACACATCAGGACAGATGGCGTATCAGATCTGCCAGTCAATAAATTTTATCA 686
DB 2426 TCATAATGACACATCAGGACAGATGGCGTATCAGATCTGCCAGTCAATAAATTTTATCA 2485
QY 687 TCATAATGACACATCAGGACAGATGGCGTATCAGATCTGCCAGTCAATAAATTTTATCA 746
DB 2486 TCATAATGACACATCAGGACAGATGGCGTATCAGATCTGCCAGTCAATAAATTTTATCA 2545
QY 747 CTATGTAGTCAAAATATTAATCTTAAATGAAACGTGATGAAACGCACTGAAACACACAGA 806
DB 2546 CTATGTAGTCAAAATATTAATCTTAAATGAAACGTGATGAAACGCACTGAAACACACAGA 2605
QY 807 TATCAAAAGTTGTTGACCTTTGGAAGTGAACGATGATGATGATGATGATGATGATGATGATGAT 866
DB 2606 TATCAAAAGTTGTTGACCTTTGGAAGTGAACGATGATGATGATGATGATGATGATGATGATGAT 2665
QY 867 GTCTACCCGGCAGCTACAGAGCTCCGAGGTCAATTTGGCTTTAGGTGCTCTCAGCCCTTG 926
DB 2666 GTCTACCCGGCAGCTACAGAGCTCCGAGGTCAATTTGGCTTTAGGTGCTCTCAGCCCTTG 2725
QY 927 TGATGTTTGGAGCATAGGTTGCAATTTCTTATTAATGATTAATGATTAATGATTAATGATTAAT 986
DB 2726 TGATGTTTGGAGCATAGGTTGCAATTTCTTATTAATGATTAATGATTAATGATTAATGATTAAT 2785
QY 987 GACTCATGATAGTAAAGAGCACCTGGCAATGATGAAACGAAATATTAGAACCCATACACACA 1046
DB 2786 GACTCATGATAGTAAAGAGCACCTGGCAATGATGAAACGAAATATTAGAACCCATACACACA 2845
QY 1047 ACACATGATTCAGAAAAACAAAGAACGCAAGTATTTTCCACCAACACAGCTAGATTGGGA 1106
DB 2846 ACACATGATTCAGAAAAACAAAGAACGCAAGTATTTTCCACCAACACAGCTAGATTGGGA 2905
QY 1107 TGAACACAGATTCTGCTGGTAGATATGTTAGGAGAGCGCTGCAAAACCGTTTGAAGGAATTTAT 1166
DB 2906 TGAACACAGATTCTGCTGGTAGATATGTTAGGAGAGCGCTGCAAAACCGTTTGAAGGAATTTAT 2965
QY 1167 GCTTTGTCATGATGAAGAACATGAGAAACCTGTTGACCTGGTTCGAAGAAATTTAGAAATA 1226
DB 2966 GCTTTGTCATGATGAAGAACATGAGAAACCTGTTGACCTGGTTCGAAGAAATTTAGAAATA 3025
QY 1227 TGATCCCAACTCAAGAAATTTACCTTGGATGAGAGCTGAGCATCTCTTTTTCAGCTTATT 1286
DB 3026 TGATCCCAACTCAAGAAATTTACCTTGGATGAGAGCTGAGCATCTCTTTTTCAGCTTATT 3085
QY 1287 AAAAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATATCTCTCTAGAGAGAGATTACT 1346
DB 3086 AAAAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATATCTCTCTAGAGAGAGATTACT 3145
QY 1347 TAAGACTGTGTGAGTCAACTAAACATTTCTTAATATTTTGTGTAACATTAATAATTTTGTGA 1406
DB 3146 TAAGACTGTGTGAGTCAACTAAACATTTCTTAATATTTTGTGTAACATTAATAATTTTGTGA 3205
QY 1407 CAGTTAACTGTAAATATTTGATGTTTGTATCAATAGCATTAATTAATTTTGTGTAACAGT 1466
DB 3206 CAGTTAACTGTAAATATTTGATGTTTGTATCAATAGCATTAATTAATTTTGTGTAACAGT 3265
QY 1467 ATGCTCTTGATTAATGCAATTAAGAAAAATTAATAATTTTCTTTTGTGAAATTTACCAATT 1526
DB 3266 ATGCTCTTGATTAATGCAATTAAGAAAAATTAATAATTTTCTTTTGTGAAATTTACCAATT 3325
QY 1527 TTAATACCTTTGAAATATCTTTTGTGTCAGTGAATAATGATGATGATGATGATGATGATGATGAT 1586
DB 3326 TTAATACCTTTGAAATATCTTTTGTGTCAGTGAATAATGATGATGATGATGATGATGATGATGAT 3385
QY 1587 TACATGAGGTCACCTCTGAAGTCAATTTTGTGTAACAAAGAAATTTTGTGTAACAAAGT 1646
DB 3386 TACATGAGGTCACCTCTGAAGTCAATTTTGTGTAACAAAGAAATTTTGTGTAACAAAGT 3445

Qy	1647	TATTCCTTAAAGGAATATCTTTTATATACCTCAAAATTTAGAACTTAACTTTAAAGTGT	1706
Ds	3446	TATTCCTTAAAGGAATATCTTTTATATACCTCAAAATTTAGAACTTAACTTTAAAGTGT	3505
Qy	1707	CTTCTGTAATTTGTTGAACGGGTGATTATTAATACTCTAGATAAGCAGGTCTAGAAACC	1766
Ds	3506	CTTCTGTAATTTGTTGAACGGGTGATTATTAATACTCTAGATAAGCAGGTCTAGAAACC	3565
Qy	1767	AAAACTCAGAAAATGTTTACTGTTAGAAATCTCTATTAATAATTTTAAAGTGTGTTCTTTT	1826
Ds	3566	AAAACTCAGAAAATGTTTACTGTTAGAAATCTCTATTAATAATTTTAAAGTGTGTTCTTTT	3625
Qy	1827	CATTGGGTGATGTCAGGGTGATTAACACAGATTTCAATGAAAGGCGATGCGATTTGTCATT	1886
Ds	3626	CATTGGGTGATGTCAGGGTGATTAACACAGATTTCAATGAAAGGCGATGCGATTTGTCATT	3685
Qy	1887	GTGACAGTTGTTTAAATAAACACATACACACTTTATTAAGATTAATACTAACTGGA	1946
Ds	3686	GTGACAGTTGTTTAAATAAACACATACACACTTTATTAAGATTAATACTAACTGGA	3745
Qy	1947	AAGTCAGCTTGGAAAATGCAATTTCCAAAGTATGTTTGGTGAGTCACAGATATAAAATA	2006
Ds	3746	AAGTCAGCTTGGAAAATGCAATTTCCAAAGTATGTTTGGTGAGTCACAGATATAAAATA	3805
Qy	2007	GAAATTCGTAGCAGAGGTTTCAGTTTTTAATACCAAGTCTTTAGGAGTCTTAAACATGGC	2066
Ds	3806	GAAATTCGTAGCAGAGGTTTCAGTTTTTAATACCAAGTCTTTAGGAGTCTTAAACATGGC	3865
Qy	2067	CAGCATCTGTTTATCAAAATGACATAAATACGTAACCTTAAAGATTAAGTTTATTAAT	2126
Ds	3866	CAGCATCTGTTTATCAAAATGACATAAATACGTAACCTTAAAGATTAAGTTTATTAAT	3925
Qy	2127	AGGCAATTTATGTCGTGATTAATCTTACGGGAGAAAGAGGATTTGATTGGAAGCAGTT	2186
Ds	3926	AGGCAATTTATGTCGTGATTAATCTTACGGGAGAAAGAGGATTTGATTGGAAGCAGTT	3985
Qy	2187	TGGGAAGAAAGTCTGCTGAAATTTCCAGAAATTTAAATGATGGTTACAT	2236
Ds	3986	TGGGAAGAAAGTCTGCTGAAATTTCCAGAAATTTAAATGATGGTTACAT	4035
RESULT 9			
PCT-US01-14827-2154			
; Sequence 2154, Application PC/TUS0114827			
; GENERAL INFORMATION:			
; APPLICANT: Hyseq, Inc			
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES			
; FILE REFERENCE: 21272-104			
; CURRENT APPLICATION NUMBER: PCT/US01/14827			
; PRIOR FILING DATE: 2001-05-16			
; PRIOR APPLICATION NUMBER: 09/577,408			
; PRIOR FILING DATE: 2000-05-18			
; NUMBER OF SEQ ID NOS: 16102			
; SOFTWARE: Custom			
; SEQ ID NO 2154			
; LENGTH: 2494			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: SIMILAR			
; LOCATION: (81)..(203)			
; OTHER INFORMATION: 92% homologous to Homo sapiens Novel protein kinase, SEQ ID			
; OTHER INFORMATION: NO: 175, accession number AAB5648, Smith-Waterman Score=213.			
PCT-US01-14827-2154			
Query Match 80.9%; Score 1825; DB 1; Length 2494;			
Best Local Similarity 92.4%; Pred. No. 0;			
Matches 2093; Conservative 0; Mismatches 60; Indels 112; Gaps 12;			
Qy	14	GTCAATTTTAAAGCAAGCTCTTGAATGACGAGATTTATCGGACCGGAGATACGTTG	73
Ds	309	GTCAATTTTAAAGCAAGCTCTTGAATGACGAGATTTATCGGACCGGAGATACGTTG	368

Qy	74	ACGAATA CAGGAATGACTACTGTGAAGGATATGTTCTTAGACATATATCACAGACATTTG	133
Ds	369	ACGAATA CAGGAATGACTACTGTGAAGGATATGTTCTTAGACATATATCACAGACATTTG	428
Qy	134	AAACGGGTATCGAATCCACTGCACTGCAATCTTTCAGTCGCGAGGAGAGCAGTCTCTA	193
Ds	429	AAACGGGTATCGAATCCACTGCACTGCAATCTTTCAGTCGCGAGGAGAGCAGTCTCTA	488
Qy	194	AAAGGAAGCGCAATAGACACATGTTTCAAGTCATCAGTCACGTTTCCGAAGCACCACGAGGA	253
Ds	489	AAAGGAAGCGCAATAGACACATGTTTCAAGTCATCAGTCACGTTTCCGAAGCACCACGAGGA	536
Qy	254	AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTTCACCTGATCTGTCAAAAGTGAGAGCG	313
Ds	537	AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTTCACCTGATCTGTCAAAAGTGAGAGCG	536
Qy	314	TTCTTAAGAGCAAGATATGNAATCGTGGACACTTTTGGGTGAAGGAGCCTTTTGGCAAGTTG	373
Ds	537	TTCTTAAGAGCAAGATATGNAATCGTGGACACTTTTGGGTGAAGGAGCCTTTTGGCAAGTTG	578
Qy	374	TAGAGTGCATTTGATCATGGCATGCGATGCGCATGTCATGTAGCAGTGAAATTCGTAAATAATG	433
Ds	579	TAGAGTGCATTTGATCATGGCATGCGATGCGCATGTCATGTAGCAGTGAAATTCGTAAATAATG	638
Qy	434	TAGGCGGTACCGTGAAGCAGCTCGTTCCAGAAATCCAAAGTATTTAGAGCAGCTTAAATAGTA	493
Ds	639	TAGGCGGTACCGTGAAGCAGCTCGTTCCAGAAATCCAAAGTATTTAGAGCAGCTTAAATAGTA	698
Qy	494	CTGATCCCAATAGTGTCTTCGAGATGTCGAGATGCTAGAAATGTTTTCATCATCATGTC	553
Ds	699	CTGATCCCAATAGTGTCTTCGAGATGTCGAGATGCTAGAAATGTTTTCATCATCATGTC	758
Qy	554	ATGTTTGTATTTGTTGAACTACTTGGGACTTTAGTACTTACGATTTCAATTAAGAAAAACA	613
Ds	759	ATGTTTGTATTTGTTGAACTACTTGGGACTTTAGTACTTACGATTTCAATTAAGAAAAACA	818
Qy	614	GCTTTCTGCCATTTCAAATTTGACCAATCAGATCTGAAGCTCTGAAGCTCTGAAGCTCTGA	673
Ds	819	GCTTTCTGCCATTTCAAATTTGACCAATCAGATCTGAAGCTCTGAAGCTCTGAAGCTCTGA	878
Qy	674	TAAATTTTATACATCAATAATAATTAACCCATACAGATCTGAAGCTCTGAAGCTCTGAAG	733
Ds	879	TAAATTTTATACATCAATAATAATTAACCCATACAGATCTGAAGCTCTGAAGCTCTGAAG	938
Qy	734	TTGTGAAGTCTGACTATGCTCAAAATATAATTTCTAAATGAAACGCTGATGAACGACAC	793
Ds	939	TTGTGAAGTCTGACTATGCTCAAAATATAATTTCTAAATGAAACGCTGATGAACGACAC	998
Qy	794	TGAAAAACACAGATATCAAAGTGTGACTTTTGGAGTGCACCGTATGATGATGAACATC	853
Ds	999	TGAAAAACACAGATATCAAAGTGTGACTTTTGGAGTGCACCGTATGATGATGAACATC	1058
Qy	854	ACAGTACTTTGTGTCTACCCGCACTACAGAGTCTCCGAGTCAATTTTGGCTTTAGGTT	913
Ds	1059	ACAGTACTTTGTGTCTACCCGCACTACAGAGTCTCCGAGTCAATTTTGGCTTTAGGTT	1118
Qy	914	GGTCTCAGCCTTTGATGTTTGGAGCATAGTGTGCTTCTTATTTGAAATTA--CCTTGG	971
Ds	1119	GGTCTCAGCCTTTGATGTTTGGAGCATAGTGTGCTTCTTATTTGAAATTA--CCTTGG	1178
Qy	972	TTTCAAGTCTTTTCAGAC--TCATGATAGTAAAGAGCCTCGCAATGATGAACGAAAT	1030
Ds	1179	TTTCAAGTCTTTTCAGAC--TCATGATAGTAAAGAGCCTCGCAATGATGAACGAAAT	1238
Qy	1031	TAGGACCCATA---CCACAAACACATGATT--CAGAAAAACAAGAAACGCAAGTATTTTCCAC	1086
Ds	1239	TAGGACCCATA---CCACAAACACATGATT--CAGAAAAACAAGAAACGCAAGTATTTTCCAC	1298
Qy	1087	C---ATAACACAGCTAGATTGGATGAAACAGGTTTC--TGCTGGTATGATATGTTAGGAGAC	1141
Ds	1299	CCATTAACCCAGCCTAGATTGGATGAAACAGGTTTC--TGCTGGTATGATATGTTAGGAGAC	1358
Qy	1142	GCTGCAAAACCGTTGAAGGAATTTATGCTTGTCTGATGAAGAACATGAGAAACATGTTTG	1201

[illegible]


```
Qy 614 GCTTTCTGCCATTTCAAAATTCACACATCAGCAGATGGCGTATCAGATCTGCCAGTCAA 673
Db 819 GCTTTCTGCCATTTCAAAATTCACACATCAGCAGATGGCGTATCAGATCTGCCAGTCAA 878
Qy 674 TAAATTTTTTACATCATATAAATTAACCCATACAGATCTGAAGCCCTGAAATATTTTGT 733
Db 879 TAAATTTTTTACATCATATAAATTAACCCATACAGATCTGAAGCCCTGAAATATTTTGT 938
Qy 734 TTGTGAAGTCTGACATAGTAGTCAAAATATAATTTCAAAATGAAGAGTGAAGCGCAC 793
Db 939 TTGTGAAGTCTGACATAGTAGTCAAAATATAATTTCAAAATGAAGAGTGAAGCGCAC 998
Qy 794 TGAATAACACAGATATCAAAAGTGTGTGACTTTTGAAGTGAAGCTGATGATCAATC 853
Db 999 TGAATAACACAGATATCAAAAGTGTGTGACTTTTGAAGTGAAGCTGATGATCAATC 1058
Qy 854 ACAGTACTTTGGTGTCTACCCGGCCTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGGTT 913
Db 1059 ACAGTACTTTGGTGTCTACCCGGCCTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGGTT 1118
Qy 914 GGTCTCAGCCTTTGATGTTTGGAGCATAGTGTGATCTTATTGAAATATTA--CCTTGG 971
Db 1119 GGTCTCAGCCTTTGATGTTTGGAGCATAGTGTGATCTTATTGAAATATTAACCCCTGG 1178
Qy 972 TTTTCACAGTCTTTCAGAC--TCATGATAGTAAAGAGCCTCGCAATGATGGAAGCAATAT 1030
Db 1179 GTTTCACAGTCTTTCAGACCTTCAATGATAGAAAGAGCCTCGCAATGATGGAAGCAATAT 1238
Qy 1031 TAGGACCCCAT--CCACAAACACATGATT--CAGAAAAAAGAAACGCAAGTATTTTTCAC 1086
Db 1239 TTAGGACCCCATACCCACAAACATGATTCCAGAAAAAAGAAAGCGCAAGTATTTTTCAC 1298
Qy 1087 C---ATAACAGCTAGATTGGAGTGAACACAGATTC--TGCTGGTAGATATGTTAGGAGAC 1141
Db 1299 CCATAACCCAGCCTAGATTGGAGTGAACACAGATTCCTGCTGGTAGATATGTTAGGAGAC 1358
Qy 1142 GCTGCAAAACCGTTGAAGAAATTTATGCTTCTCATGATGAAGAAACATGAGAAATCTGTTG 1201
Db 1359 GCTGCAAAACCGTTGAAGAAATTTATGCTTCTCATGATGAAGAAACATGAGAAATCTGTTG 1418
Qy 1202 ACCTGGTTGGAAGAAATTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCAT 1261
Db 1419 ACCTGGTTGGAAGAAATTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCAT 1478
Qy 1262 TGCAGCATCTTTCTTTGACATTTTAAAGAAATGAAATGGGAATCAGTGTCTTACT 1321
Db 1479 TGCAGCATCTTTCTTTGACATTTTAAAGAAATGAAATGGGAATCAGGAGTCTTACT 1538
Qy 1322 ATATCTCTCTAGAGAGATTTACTTAAGACTGTGTCACTCAACTCAAACTTAAATATTT 1381
Db 1539 ATATCTCTCTAGAGAGATTTACTTAAGACTGTGTCACTCAACTCAAACTTAAATATTT 1598
Qy 1382 TTTGTPAAACATTTAAATTTTGTACAGTTAAGTGAATATTTGATGTTTGTATCAAT 1441
Db 1599 TTTGTPAAACATTTAAATTTTGTACAGTTAAGTGAATATTTGATGTTTGTATCAAT 1658
Qy 1442 AGCATAATTAATCTTTGAAGCAAGTATGCTTGTGATGATGATGATGATGATGATGAT 1501
Db 1659 AGCATAATTAATCTTTGAAGCAAGTATGCTTGTGATGATGATGATGATGATGATGAT 1718
Qy 1502 ATTTTCTTTTGAATTTTAAATTTTAAATACCTTTTGAATATCTTTTGTGTCAGTGA 1561
Db 1719 ATTTTCTTTTGAATTTTAAATTTTAAATACCTTTTGAATATCTTTTGTGTCAGTGA 1778
Qy 1562 TAAATGTGATGATCTTGCCTTTTGTACATGGAGTCACTCTGAAGTGAATTTTGTGA 1621
Db 1779 TAAATGTGATGATCTTGCCTTTTGTACATGGAGTCACTCTGAAGTGAATTTTGTGA 1838
Qy 1622 GTAAAGGAAATCTTGACTTCTTTATTTCTTAAAGGAATATTTCTTTATATCTTCAAT 1681
Db 1839 GTAAAGGAAATCTTGACTTCTTTATTTCTTAAAGGAATATTTCTTTATATCTTCAAT 1898
```

```
Qy 1682 TTAGAACTTAACCTTTAAAGATTTTCTCTGTAATTTGTTGAACGGGTGATATTAATTAAC 1741
Db 1899 TTAGAACTTAACCTTTAAAGATTTTCTCTGTAATTTGTTGAACGGGTGATATTAATTAAC 1958
Qy 1742 TCTAGATAAGCAGTACTAGAAACCAAACTCAGAAATGTTTACTGTGTAGAAATCTTATT 1801
Db 1959 TCTAGATAAGCAGTACTAGAAACCAAACTCAGAAATGTTTACTGTGTAGAAATCTTATT 2018
Qy 1802 AAATTTTAAAGTGTGTATTTCTTTTCAATTTGGGTGATGTCAGGGTGTAAACC--AGACATT 1859
Db 2019 AAATTTTAAAGTGTGTATTTCTTTTCAATTTGGGTGATGTCAGGGTGTAAACCAGACATT 2078
Qy 1860 CATGGAAGG---CATGCACTTTGTCCATTTGTGCACAGTTTGTGTTTAAATAAACCCAC--ATAC 1915
Db 2079 CATGGAAGGCGCATGCCAGTTTGTGCCATTTGTGCACAGTTTGTGTTTAAATAAACCCACGATAC 2138
Qy 1916 ACATTTTATTAAGATTAATAATC--TAACTGGAAGTCACTTGGAAAAATGGAAATTTCC 1973
Db 2139 ACATTTTATTAAGATTTAAATCGTAATCTGGAAGGTCACTTGGAAAAATGGAAATTTCC 2198
Qy 1974 AAGTATGCTTTGGTGAGTCAAGATATAAAAAATAGAAATTTCTGATGAGAGTTTTCAGTTTT 2033
Db 2199 AAGTATGCTTTGGTGAGTCAAGATATAAAAAATAGAAATTTCTGATGAGAGTTTTCAGTTTT 2258
Qy 2034 TAATACCAAGTCTCTTAGGAGTCTTAAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAA 2093
Db 2259 TAATACCAAGTCTCTTAGGAGTCTTAAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAA 2318
Qy 2094 TAGTAACCTATAAGAAATTT--AAGTTTATTAATTAGGCAATTTATGCTGTGATTAATTC 2151
Db 2319 TAGTAACCTATAAGAAATTTAAGGTTTATTAATTAGGCAATTTATGCTGTGATTAATTC 2378
Qy 2152 TTACGGGAGAAAGAGGATTTGATTGGAAGCAGTTTGGGAAGAAAGTCTGCTGAAATTT 2211
Db 2379 TTACGGGAGAAAGAGGATTTGATTGGAAGCAGTTTGGGAAGAAAGTCTGCTGAAATTT 2438
Qy 2212 CCAGAAATTTAATGATTTGTTTACATAAACTTTTGTGACTTCAGAAA 2256
Db 2439 CCAGAAATTTAATGATTTGTTTACATAAACTTTTGTGCTTCAAAA 2483
```

```
RESULT 11
US-10-170-235-20526
; Sequence 20526, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
; FILE REFERENCE: C1001380
; CURRENT APPLICATION NUMBER: US/10/170,235
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 20526
; LENGTH: 1795
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-20526
```

```
Query Match 72.58; Score 1635; DB 49; Length 1795;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 14 GTCAATTTTAGAAGCAAGGTCTTTGAATGAGCGAGATATTCGGGACCGGAGATACGTTG 73
Db 161 GTCAATTTTAGAAGCAAGGTCTTTGAATGAGCGAGATATTCGGGACCGGAGATACGTTG 220
Qy 74 ACGAATACAGAATGACTCTGTGAAGGATATGTTCTTAGACATTTATCAGAGACATTTG 133
Db 221 ACGAATACAGAATGACTCTGTGAAGGATATGTTCTTAGACATTTATCAGAGACATTTG 280
Qy 134 AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGAGCAGGAGAGCAGTCTCTA 193
Db 281 AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGAGCAGGAGAGCAGTCTCTA 340
```

194 AAAGGAGCGCAATAGACACTGTTCAAGTCATCATGTCACGTTTCGAGAGCCACCGAAGGA 253
Db
341 AAAGGAGCGCAATAGACACTGTTCAAGTCATCATGTCACGTTTCGAGAGCCACCGAAGGA 400
Qy
254 AAAGATCCAGGATATAGAGGATGATGAGGAGGTCACCTGATCTCTCAAGTGGAGAGG 313
Db
401 AAAGATCCAGGATATAGAGGATGATGAGGAGGTCACCTGATCTCTCAAGTGGAGAGG 460
Qy
314 TTCTAAGAGCAGATATGAAATCGTGGACACATCTTGGGTGAAGGAGCCCTTGGCAAGTTG 373
Db
461 TTCTAAGAGCAGATATGAAATCGTGGACACATCTTGGGTGAAGGAGCCCTTGGCAAGTTG 520
Qy
374 TAGAGTGCATTGATCATGGCATGGATGGATGCATGTAGCAGTGAAGATCGTAAAGAAATG 433
Db
521 TAGAGTGCATTGATCATGGCATGGATGGATGCATGTAGCAGTGAAGATCGTAAAGAAATG 580
Qy
434 TAGGCGGTACCGTGAAGCAGCTCGTTCCAGATCTGATGATGATGATGATGATGATGATG 493
Db
581 TAGGCGGTACCGTGAAGCAGCTCGTTCCAGATCTGATGATGATGATGATGATGATGATG 640
Qy
494 CTGATCCCAATAGTGTCTTCCGATGCTCCAGATGCTAGATGCTGATGATGATGATGATG 553
Db
641 CTGATCCCAATAGTGTCTTCCGATGCTCCAGATGCTAGATGCTGATGATGATGATGATG 700
Qy
554 ATGTTTGTATTGTTGAACTACTGGGACTTAGTACTTACGATTTCAATTAAGAAAAACA 613
Db
701 ATGTTTGTATTGTTGAACTACTGGGACTTAGTACTTACGATTTCAATTAAGAAAAACA 760
Qy
614 GCTTCTGCGCATTTCAAATGACCAATCAGCAGATGCGGATGATGATGATGATGATGATG 673
Db
761 GCTTCTGCGCATTTCAAATGACCAATCAGCAGATGCGGATGATGATGATGATGATGATG 820
Qy
674 TAAATTTTTCATCATATAATTAATTAACCATACAGATCTGAAGCTGAAATATTTTGT 733
Db
821 TAAATTTTTCATCATATAATTAATTAACCATACAGATCTGAAGCTGAAATATTTTGT 880
Qy
734 TTGGAAGTCTGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 793
Db
881 TTGGAAGTCTGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 940
Qy
794 TGAAGAACACAGATATCAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 853
Db
941 TGAAGAACACAGATATCAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1000
Qy
854 ACAGTACTTTGTTGTTCTACCGGCACTACAGATCTCCGAGTCAATTTTGGCTTTAGTT 913
Db
1001 ACAGTACTTTGTTGTTCTACCGGCACTACAGATCTCCGAGTCAATTTTGGCTTTAGTT 1060
Qy
914 GGTCTCAGCCTTTGATGTTTGGAGCATAGTTGATTTCTTATTTGAATATTTACCTTTGTT 973
Db
1061 GGTCTCAGCCTTTGATGTTTGGAGCATAGTTGATTTCTTATTTGAATATTTACCTTTGTT 1120
Qy
974 TCAGAGTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1033
Db
1121 TCAGAGTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1180
Qy
1034 GACCCATACCAACACATGATTTAGAGAAACAGAGAAACAGAGATTTTTCACCATACCC 1093
Db
1181 GACCCATACCAACACATGATTTAGAGAAACAGAGAAACAGAGATTTTTCACCATACCC 1240
Qy
1094 AGCTAGATTTGGATGAACACATGTTCTGCTGGTAGATATGTTAGGAGAGCGCTGCAACCGT 1153
Db
1241 AGCTAGATTTGGATGAACACATGTTCTGCTGGTAGATATGTTAGGAGAGCGCTGCAACCGT 1300
Qy
1154 TGAAGGAATTTATGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1213
Db
1301 TGAAGGAATTTATGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1360
Qy
1214 GAATGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAGCATTTGAGCATCTCT 1273
Db
1361 GAATGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAGCATTTGAGCATCTCT 1420

1274 TCTTTGACTTATTAATAAAGAAATGAAATGGAATCAAGTGGTCTTATCTATATATCTTCT 1333
Db
1421 TCTTTGACTTATTAATAAAGAAATGAAATGGAATCAAGTGGTCTTATCTATATATCTTCT 1480
Qy
1334 AGAGAGATTAATTAAGACTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATT 1393
Db
1481 AGAGAGATTAATTAAGACTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATT 1540
Qy
1394 AAATTAATTTTGTACAGTTAAGTGAATATTTGTATTTTGTATCAATAGCATATTAAC 1453
Db
1541 AAATTAATTTTGTACAGTTAAGTGAATATTTGTATTTTGTATCAATAGCATATTAAC 1600
Qy
1454 TTGTTAAGCAAGTATGTCCTTGAATATCAATAGCAAAATTAATAATTTTCTTTT 1513
Db
1601 TTGTTAAGCAAGTATGTCCTTGAATATCAATAGCAAAATTAATAATTTTCTTTT 1660
Qy
1514 GAATTAACCAATTTTAATACCTTTGAATATCAATAGCAAAATTTTGTATCAATAGCAATG 1573
Db
1661 GAATTAACCAATTTTAATACCTTTGAATATCAATAGCAAAATTTTGTATCAATAGCAATG 1720
Qy
1574 ATCTTGCTTTTGTACAGTGGTCACTCTGAACTGATTTTGTATCAATAGCAAAAT 1633
Db
1721 ATCTTGCTTTTGTACAGTGGTCACTCTGAACTGATTTTGTATCAATAGCAAAAT 1780
Qy
1634 CTGACTACTTTATA 1648
Db
1781 CTGACTACTTTATA 1795

RESULT 12

US-10-932-349-580

; Sequence 580, Application US/10932349

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele

; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES

; TITLE OF INVENTION: ENCODING HUMAN KINASE PROTEINS, METHODS OF DETECTION AND

; TITLE OF INVENTION: USES THEREOF

; FILE REFERENCE: CL001483

; CURRENT APPLICATION NUMBER: US/10/932,349

; CURRENT FILING DATE: 2004-09-02

; NUMBER OF SEQ ID NOS: 123188

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 580

; LENGTH: 1795

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-932-349-580

Query Match 72.5%; Score 1634.6; DB 65; Length 1795;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1634; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 14 GTCATTATTTAGACGAGTCTTGAATGACGAGATTTATCGGACCGGAGATACGTTG 73
Db 161 GTCATTATTTAGACGAGTCTTGAATGACGAGATTTATCGGACCGGAGATACGTTG 220
Qy 74 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTTATCAGAGACATTTG 133
Db 221 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTTATCAGAGACATTTG 280
Qy 134 AAAGCGGTATCGAATCCATCGCAGTAAATCTTCAGTCCGAGCAGGAGAGAGCAGTCTTA 193
Db 281 AAAGCGGTATCGAATCCATCGCAGTAAATCTTCAGTCCGAGCAGGAGAGAGCAGTCTTA 340
Qy 194 AAAGGAGCGCAATAGACACTGTTCAAGTCATCATGTCAGTTCGAGAGCCACCGAAGGA 253
Db 341 AAAGGAGCGCAATAGACACTGTTCAAGTCATCATGTCAGTTCGAGAGCCACCGAAGGA 400
Qy 254 AAAGATCCAGGATATGAGGATGATGAGGAGGTCACCTGATCTCTGTCAGAGTGGAGAGG 313
Db 401 AAAGATCCAGGATATGAGGATGATGAGGAGGTCACCTGATCTCTGTCAGAGTGGAGAGG 460
Qy 314 TTCTAAGAGCAAGATATGAAATCGTGGACACTTTTGGGTGAAGGAGCCCTTTGCAAAAGTTG 373

Qy	494	CTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGCTAGAGTGTGTGATCATCATGGTC	553
Db	641	CTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGCTAGAGTGTGTGATCATCATGGTC	700
Qy	554	ATGTTTGTATTGTCTTTGAACTACTCGGGACTTAGTACTTACGATTTTCATTAAGAAAAACA	613
Db	701	ATGTTTGTATTGTCTTTGAACTACTCGGGACTTAGTACTTACGATTTTCATTAAGAAAAACA	760
Qy	614	GCCTTCTGCCATTTCAAATGACCACATCAGGCGAGATGGCGTATCAGATCTGCCAGTCAA	673
Db	761	GCCTTCTGCCATTTCAAATGACCACATCAGGCGAGATGGCGTATCAGATCTGCCAGTCAA	820
Qy	674	TAAATTTTTTACATCAATAATAAATTAAACCCATACAGATCTGAAGCCCTGAAAAATATTTGT	733
Db	821	TAAATTTTTTACATCAATAATAAATTAAACCCATACAGATCTGAAGCCCTGAAAAATATTTGT	880
Qy	734	TTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTAAATGAAACGCTGATGAAGCGCAC	793
Db	881	TTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTAAATGAAACGCTGATGAAGCGCAC	940
Qy	794	TGAAAAACACAGATATCAAAGTCTGTCACTTTTGGAACTGCAACGTATGATGAACATC	853
Db	941	TGAAAAACACAGATATCAAAGTCTGTCACTTTTGGAACTGCAACGTATGATGAACATC	1000
Qy	854	ACAGTACTTTGGTGTCTACCCGGCACTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGGTT	913
Db	1001	ACAGTACTTTGGTGTCTACCCGGCACTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGGTT	1060
Qy	914	GGTCTCAGCCTTGTGATGTTTGGAGCATAGTTTGCATCTTTTGAATATATACCTTGGTT	973
Db	1061	GGTCTCAGCCTTGTGATGTTTGGAGCATAGTTTGCATCTTTTGAATATATACCTTGGTT	1120
Qy	974	TCAGAGCTTTTCAGACTCATGATAGTAAAGAGCACCTGGCAATCATGGAAGCAATATTAG	1033
Db	1121	TCAGAGCTTTTCAGACTCATGATAGTAAAGAGCACCTGGCAATCATGGAAGCAATATTAG	1180
Qy	1034	GACCCATACCAACACATGATTCAGAAAAACAGAAAAACGCAAGTATTTTCCACCATAACC	1093
Db	1181	GACCCATACCAACACATGATTCAGAAAAACAGAAAAACGCAAGTATTTTCCACCATAACC	1240
Qy	1094	AGCTAGATTGGGATGAACACAGTTCTCGCTGTGTAGATAGTTAGGAGACGCTGCAAAACCGT	1153
Db	1241	AGCTAGATTGGGATGAACACAGTTCTCGCTGTGTAGATAGTTAGGAGACGCTGCAAAACCGT	1300
Qy	1154	TGAAGGAATTTATGCTTTTGCATCATGAGAACATGAGAACTGTTTGCCTGGTTCGAA	1213
Db	1301	TGAAGGAATTTATGCTTTTGCATCATGAGAACATGAGAACTGTTTGCCTGGTTCGAA	1360
Qy	1214	GAATGTTAGAAATATGATCCAACTCAAAGAATTACCTTTGGATGAAGCATTGCAGCATCCTT	1273
Db	1361	GAATGTTAGAAATATGATCCAACTCAAAGAATTACCTTTGGATGAAGCATTGCAGCATCCTT	1420
Qy	1274	TCCTTGTACTTATTAATAAGAAATGAAATGGGAATCAAGTGGTCTTACTATATCTCTCT	1333
Db	1421	TCCTTGTACTTATTAATAAGAAATGAAATGGGAATCAAGTGGTCTTACTATATCTCTCT	1480
Qy	1334	AGAGAGACTTACTTAAAGACTGTGCAGTCAACTAAACATTTCTATAATTTTGTGAACATT	1393
Db	1481	AGAGAGACTTACTTAAAGACTGTGCAGTCAACTAAACATTTCTATAATTTTGTGAACATT	1540
Qy	1394	AAATTTATTTTGTACAGTTAAGTGTAAATATTGTATGTTTGTATCAATAGCATAAATTAAC	1453
Db	1541	AAATTTATTTTGTACAGTTAAGTGTAAATATTGTATGTTTGTATCAATAGCATAAATTAAC	1600
Qy	1454	TTGTTAAGCAAGTATGTTCTGTAATAATGCAATAGAAAAATTAATAATTTTCTTTTTT	1513
Db	1601	TTGTTAAGCAAGTATGTTCTTGAATAATGCAATAGAAAAATTAATAATTTTCTTTTTT	1660
Qy	1514	GAAATTAACCATTTTAAATACCTTTGAAATATTCCTTTGTGCCAGTGAATAATGTGATTG	1573
Db	1661	GAAATTAACCATTTTAAATACCTTTGAAATATTCCTTTGTGCCAGTGAATAATGTGATTG	1720

QY	1574	ATCTTGCTTTTGTACATGAGGTCACCTCTGAAGTCAATTTTTTTTGAGTAAAGGAAT	1633
DB	1721	ATCTTGCTTTTGTACATGAGGTCACCTCTGAAGTCAATTTTTTTTGAGTAAAGGAAT	1780
QY	1634	CTTGACTACTTTTATA	1648
DB	1781	CTTGACTACTTTTATA	1795
RESULT 14			
US-10-756-149-49			
; Sequence 49, Application US/10756149			
; GENERAL INFORMATION:			
; APPLICANT: Aziz, Natasha			
; APPLICANT: Zlotnik, Albert			
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS OF NOVEL METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER			
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER			
; FILE REFERENCE: file			
; CURRENT APPLICATION NUMBER: US/10/756,149			
; CURRENT FILING DATE: 2004-01-12			
; NUMBER OF SEQ ID NOS: 5818			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 49			
; LENGTH: 1814			
; TYPE: DNA			
; ORGANISM: Homo Sapiens			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (1562)..(1562)			
; OTHER INFORMATION: n is a, c, g, or t			
US-10-756-149-49			
Query Match 72.3%; Score 1632; DB 62; Length 1814;			
Best Local Similarity 99.3%; Pred. No. 0;			
Matches 1638; Conservative 0; Mismatches 11; Indels 0; Gaps 0;			
QY	14	GTCATTATTAGAACGAAGTCCTTTGAATGAGCGAGATTATCGGGACCGGAGATACGTTG	73
DB	161	GTCATTATTAGAACGAAGTCCTTTGAATGAGCGAGATTATCGGGACCGGAGATACGTTG	220
QY	74	ACGAATACAGGAATGACTCTGTGAAGGATATGTTCTTAGACATTATCAGAGACATTG	133
DB	221	ACGAATACAGGAATGACTCTGTGAAGGATATGTTCTTAGACATTATCAGAGACATTG	280
QY	134	AAAGCGGGTATCGAATCCATCGCAGTAAATCTTCAGTCCGACGAGGAGAGCAGTCTTA	193
DB	281	AAAGCGGGTATCGAATCCATCGCAGTAAATCTTCAGTCCGACGAGGAGAGCAGTCTTA	340
QY	194	AAAGGAAGCGCAATAGACACTGTTTCAAGTCATCAGTCACTGTTGAAAGAGCCACCGAAGGA	253
DB	341	AAAGGAAGCGCAATAGACACTGTTTCAAGTCATCAGTCACTGTTGAAAGAGCCACCGAAGGA	400
QY	254	AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAGTGGAGACG	313
DB	401	AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAGTGGAGACG	460
QY	314	TTCTAAGACCAAGATATGAATCGTGACACTTTGGGTGAAGGAGCCTTTGGCAAAAGTTG	373
DB	461	TTCTAAGACCAAGATATGAATCGTGACACTTTGGGTGAAGGAGCCTTTGGCAAAAGTTG	520
QY	374	TAGAGTGCAATTGATCATGGCATGGCATGTCATGTAGCAGTGAATAATCGTAAAAAATG	433
DB	521	TAGAGTGCAATTGATCATGGCATGGCATGTCATGTAGCAGTGAATAATCGTAAAAAATG	580
QY	434	TAGCCGTTACCGTGAAGCAGCTCGTTTCAAGATCCAAAGTATTAGAGCATTAAATAGTA	493
DB	581	TAGCCGTTACCGTGAAGCAGCTCGTTTCAAGATCCAAAGTATTAGAGCATTAAATAGTA	640
QY	494	CTGATCCCAATAGTGTCTCCGATGTGTCAGATGCTAGAAATGCTTTGATCATCATGGTC	553
DB	641	CTGATCCCAATAGTGTCTCCGATGTGTCAGATGCTAGAAATGCTTTGATCATCATGGTC	700
QY	554	ATGTTTGTATTGTTTGAACACTACTGGGACTTAGTACTTACGATTTTCATTTAAAGAAAAACA	613

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2005, 07:34:52 ; Search time 7297.99 Seconds
(without alignments)

11766.665 Million cell updates/sec

Title: US-10-801-671-1_COPY_72_2327

Perfect score: 2256

Sequence: 1 gacactcatccagtcattt.....aaacttttgcattcagaaa 2256

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1644.4	72.9	2798	3	CR749504
2	1283.2	56.9	1556	3	BC015942 Homo sapi
3	1194	52.9	1446	9	AY408252 Homo sapi
4	1118.8	49.6	1865	3	AF212224 Homo sapi
5	1078.2	47.8	1629	3	AK013974 Mus muscu
6	978.2	43.4	1421	9	AY408254 Mus muscu
7	937.6	41.6	1446	9	AY408253 Pan trogl
8	814.8	36.1	1737	3	CR591789 full-leng
9	814.8	36.1	1746	3	CR612115 full-leng
10	814.8	36.1	1754	3	CR617472 full-leng
11	754.4	33.4	973	5	BUS16568 AGENCOURT
12	734.2	32.5	1689	3	AK076199 Mus muscu
13	722	32.0	722	4	BM786882 K-EST0065
14	720.8	32.0	966	7	CF110608 Shultzomi
15	700.8	31.1	727	1	AA631990 np74a05.s
16	700.6	31.1	729	4	BM675318 UI-E-EJO-
17	670.6	29.7	695	5	BQ774390 UI-H-E21-
18	666	28.5	1664	3	CR593610 full-leng
19	647.6	28.7	1775	3	BC028149 Homo sapi
20	640	28.4	3226	3	BC028573 Homo sapi
21	635.4	28.2	721	5	BUI184090 AGENCOURT
22	635.2	28.2	919	4	BG109862 60279328
23	633.6	28.1	659	4	BM681648 UI-E-EJO-
24	630	27.9	664	5	BM969890 UI-CF-DU1

c	25	628	27.8	1337	7	CO389096
	26	623.8	27.7	643	6	CB127989 K-EST0177
	27	621.8	27.6	849	4	BG916444 602814061
c	28	620.8	27.5	637	5	BU729717 UI-E-CK1-
	29	620.6	27.5	667	1	AJ656344 AJ656344
	30	617	27.3	1551	7	CF110517 Shultzomi
	31	611.6	27.1	957	5	BX374503 BX374503
	32	609.4	27.0	923	5	BX385193 BX385193
c	33	609	27.0	767	5	BX374502 BX374502
	34	607	26.9	607	6	CD676847 ho07d04.y
	35	594.8	26.4	1138	3	CR609176 full-leng
	36	592.2	26.2	902	1	AL540351 AL540351
	37	591.6	26.2	950	7	CF411264 CH3072.F
c	38	591	26.2	624	4	BI494554 df111a06.
	39	589.8	26.1	893	6	CB202758 AGENCOURT
	40	589.2	26.1	950	5	BP383242 BP383242
c	41	578.4	25.6	593	6	CA314357 UI-CF-FNO
	42	573.4	25.4	710	5	BQ210018 UI-R-EPO-
c	43	570.8	25.3	618	6	CB285812 CMD35.A05
	44	570	25.3	581	5	BP307701 BP307701
	45	566	25.1	568	4	BM727179 UI-E-EJO-

ALIGNMENTS

RESULT 1
CR749504
LOCUS CR749504 2798 bp mRNA linear HTC 19-AUG-2004
DEFINITION Homo sapiens mRNA; cDNA DKFZp686A20267 (from clone DKFZp686A20267).
ACCESSION CR749504
VERSION CR749504.1 GI:51476700
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2798)
AUTHORS Koehrer,K., Beyer,A., Mewes,H.W., Weill,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
CONSRM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: s.wiemann@dkz-heidelberg.de; sequenced by BMFZ (Biomedical Research Center at the Heinrich-Heine-University, Dueseldorff/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp686A20267) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering.
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp686A20267
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.

FEATURES

source
1..2798
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686A20267"
/tissue type="cerebellum"
/clone lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host DH10B; sites SfiIA + SfiIB"
/dev_stages="adult"
/note="protein serine threonine kinase Ctk4, differentially spliced"
1..2798
/gene="DKFZp686A20267"
85..246
/gene="DKFZp686A20267"
/codon_start=1
/product="hypothetical protein"

```
/protein_id="CAH18325.1"  
/db_xref="GI:51476701"  
/translation="MRHSKRTHCPDWDSRWGHSYRGSHKRRKRSHSSTQENRHCK  
PHHOFKESD"
```

ORIGIN

Query Match	72.9%	Score 1644.4;	DB 3;	Length 2798;
Best Local Similarity	99.3%	Pred. No. 0;		
Matches 1651; Conservative	0;	Mismatches 11;	Indels 0;	Gaps 0;

QY	1	GACACTCATCCAGTCATTATTAGAGCAAGTCCCTTGAATGAGCAGATTTACGGGAC	60
Db	1135	GACACTCATCCAGTCATTATTAGAGCAAGTCCCTTGAATGAGCAGATTTACGGGAC	1194
QY	61	CGGAGATACGTTGACCAATACAGAAATGACTACTGTCAAGATATGTTCTCTAGACATTAT	120
Db	1195	CGGAGATACGTTGACCAATACAGAAATGACTACTGTCAAGATATGTTCTCTAGACATTAT	1254
QY	121	CACAGAGACATTGAAACGGGTATCGAATCCACTGTCAGTAAATCTTCAGTCCGACAGG	180
Db	1255	CACAGAGACATTGAAACGGGTATCGAATCCACTGTCAGTAAATCTTCAGTCCGACAGG	1314
QY	181	AGAAGCAGTCTTAAAGAGCAAGCAATAGACATCTGTTCAAGTCATCAGTCACGTTGCAAG	240
Db	1315	AGAAGCAGTCTTAAAGAGCAAGCAATAGACATCTGTTCAAGTCATCAGTCACGTTGCAAG	1374
QY	241	AGCCACCGAAGGAAAAGATCACAGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGT	300
Db	1375	AGCCACCGAAGGAAAAGATCOAGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGT	1434
QY	301	CAAAAGTGGAGACGTTCTTAAGAGCAAGATATGAATCGTGGACACTTTGGGTGAAGAGCC	360
Db	1435	CAAAAGTGGAGACGTTCTTAAGAGCAAGATATGAATCGTGGACACTTTGGGTGAAGAGCC	1494
QY	361	TTTGGCAAAAGTTGATGATGCGATGCGCATGCGATGCGATGCGATGCGATGCGATGCGA	420
Db	1495	TTTGGCAAAAGTTGATGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGA	1554
QY	421	ATCGTAAAAAATGTPAGGCCGTTACCGTGAAGCAGCTCGTTTCAGAAATCCCAAGTATTAGAG	480
Db	1555	ATCGTAAAAAATGTPAGGCCGTTACCGTGAAGCAGCTCGTTTCAGAAATCCCAAGTATTAGAG	1614
QY	481	CACTTAAATAGTACTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGCTAGATGGTTT	540
Db	1615	CACTTAAATAGTACTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGCTAGATGGTTT	1674
QY	541	GATCATCATGGTCATGTTGTTGTTGAACTACTGGGACATTAGTACTTAGCATTTTC	600
Db	1675	GATCATCATGGTCATGTTGTTGTTGTTGAACTACTGGGACATTAGTACTTAGCATTTTC	1734
QY	601	ATTAAAGAAACACGCTTCTGCCATTTCAAAATGACCACATCAGGCAGATGGCGTATCAG	660
Db	1735	ATTAAAGAAACACGCTTCTGCCATTTCAAAATGACCACATCAGGCAGATGGCGTATCAG	1794
QY	661	ATCTGCCAGTCATAATAATTTTACATCATATAAATTAACCCATACAGATCTGAAGCCT	720
Db	1795	ATCTGCCAGTCATAATAATTTTACATCATATAAATTAACCCATACAGATCTGAAGCCT	1854
QY	721	GAAAAATATTTTGTGTGAGTCTGACTATGATGTCAAAATATAAATCTTAAAAAGAAACGT	780
Db	1855	GAAAAATATTTTGTGTGAGTCTGCGTATGTCAAAATATAAATCTTAAAAAGAAACGT	1914
QY	781	GATGAAACGACACTGAAAAACACAGATATCAAGTGTCTGACTTTTGAAGTGCAACGTAT	840
Db	1915	GATGAAACGACACTGAAAAACACAGATATCAAGTGTCTGACTTTTGAAGTGCAACGTAT	1974
QY	841	GATGATGAAACATCACAGTACTTTTGGTGTCTACCCGGCACTACAGAGTCCCGAGGTCATT	900
Db	1975	GATGATGAAACATCACAGTACTTTTGGTGTCTACCCGGCACTACAGAGTCCCGAGGTCATT	2034
QY	901	TTGGCTTTTAGGTTGGTCTCAGCCTTGTGATGTTTGGAGCATPAGGTTGCATTTATTGAA	960
Db	2035	TTGGCTTTTAGGTTGGTCTCAGCCTTGTGATGTTTGGAGCATPAGGTTGCATTTATTGAA	2094

Qy	961	TATTACCTTGGTTCACAGTCTTTTCAGACTCATGTAAAGACGACCCCTGGCAATGATG	102
Db	2095	TATTACCTTGGTTCACAGTCTTTTCAGACTCATGTAAAGACGACCCCTGGCAATGATG	2154
Qy	1021	GAACGAATATTAGGACCCATACCAACAACATGATTCAGAAAAACAAGAAACCAAGTAT	1080
Db	2155	GAACGAATATTAGGACCCATACCAACAACATGATTCAGAAAAACAAGAAACCAAGTAT	2214
Qy	1081	TTTCCACCATAAACAGCTAGATTTGGGATGAACACAGTTCTCTCGTAGATATGTTAGGAGA	1140
Db	2215	TTTCCACCATAAACAGCTAGATTTGGGATGAACACAGTTCTCTCGTAGATATGTTAGGAGA	2274
Qy	1141	CGCTGCAAAACCGTTGAAGGAATTTATGCTTTGTCTGATGAAGAAACATGAGAAATCTGTTT	1200
Db	2275	CGCTGCAAAACCGTTGAAGGAATTTATGCTTTGTCTGATGAAGAAACATGAGAAATCTGTTT	2334
Qy	1201	GACCTGGTTCGAAGAAATGTTAGAAATATGATCCAACTCAAGAAATACCTTTGGATGAAGCA	1260
Db	2335	GACCTGGTTCGAAGAAATGTTAGAAATATGATCCAACTCAAGAAATACCTTTGGATGAAGCA	2394
Qy	1261	TTGCAGCATCCTTTCCTTTGACTTATTTAAAAAGAAATGAAATCGGAAATCAGTGGTCTTAC	1320
Db	2395	TTGCAGCATCCTTTCCTTTGACTTATTTAAAAAGAAATGAAATCGGAAATCAGTGGTCTTAC	2454
Qy	1321	TATATACTTCTCTAGAAGAGATTACTTTAAGACTGTGTGTCAGTCAACTAAAAATTTCTAATAT	1380
Db	2455	TATATACTTCTCTAGAAGAGATTACTTTAAGACTGTGTGTCAGTCAACTAAAAATTTCTAATAT	2514
Qy	1381	TTTTTGTAACAATTTAAATATTTTGTACAGTTAAGTGTAAATATATGTATGTTTGTATCAA	1440
Db	2515	TTTTTGTAACAATTTAAATATTTTGTACAGTTAAGTGTAAATATATGTATGTTTGTATCAA	2574
Qy	1441	TAGCATAAATTTAACTTTGTTAAGCAAGTATGCTCTTGATAATGCATTAGAAAAATTTAAAAAT	1500
Db	2575	TAGCATAAATTTAACTTTGTTAAGCAAGTATGCTCTTGATAATGCATTAGAAAAATTTAAAAAT	2634
Qy	1501	AAATTTTCTTTTTTGAANAATACCAATTTTTTAATAACCTTTGAAAAATATCCTTTGTGTCCAGTG	1560
Db	2635	AAATTTTCTTTTTTGAANAATACCAATTTTTTAATAACCTTTGAAAAATATCCTTTGTGTCCAGTG	2694
Qy	1561	ATAAATGTGATTCGATCTTGTGCTTTTGTGTACATGAGGTCACCTCTGAAAGTCATTTTTTTTTG	1620
Db	2695	ATAAATGTGATTCGATCTTGTGCTTTTGTGTACATGAGGTCACCTCTGAAAGTCATTTTTTTTTG	2754
Qy	1621	AGTAAAAAGGAAATCTTGACTACTTTATTTATTTCTTTAAAGGAATA	1662
Db	2755	AGTAAAAAGGAAATCTTGACTACTTTTAAAAAATAAAAAAAAAAAAAA	2796

RESULT	2
BC015942	
LOCUS	Homo sapiens CDC-like kinase 4, mRNA (cdna clone IMAGE:3908500), linear HTC 19-NOV-2003
DEFINITION	containing frame-shift errors.
ACCESSION	BC015942
VERSION	BC015942.1 GI:16198521
KEYWORDS	HTC.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1556) Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefcr,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J.J., Hsieh,F., Diatchenko,L., Marsina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carinci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullihy,S.U., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.
AUTHORS	

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalilus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 1556)

Strausberg, R.

Direct Submission

Submitted (15-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgaps-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalilus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAX Plate: 15 Row: d Column: 16

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 10190705

This clone has the following problem: frame shifted.

Location/Qualifiers

1..1556

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3908500"

/tissue_type="Uterus, leiomyosarcoma"

/clone_lib="NIH MGC_71"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

ORIGIN

Query Match 56.9%; Score 1283.2; DB 3; Length 1556;

Best Local Similarity 99.8%; Pred. No. 2e-265;

Matches 1285; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 14 GTCATTATTAGAACCAAGGCTCTTGAATGACGAGATTATCGGACCGGAGATACGTTG 73

Db 269 GTCATTATTAGAACCAAGGCTCTTGAATGACGAGATTATCGGACCGGAGATACGTTG 328

Qy 74 ACGAATACAGGAATGACTACTGTGAGGATATGTTCTTAGACATTATCACAGACATTG 133

Db 329 ACGAATACAGGAATGACTACTGTGAGGATATGTTCTTAGACATTATCACAGACATTG 388

Qy 134 AAAGCGGTATCGAATCCACTGCAGTAATCTTCAGTCGCGAGCAGGAGAGCAGTCCTA 193

Db 389 AAAGCGGTATCGAATCCACTGCAGTAATCTTCAGTCGCGAGCAGGAGAGCAGTCCTA 448

Qy 194 AAAGGAAGCGCAATAGACACCTGTTCAAGTCATCAGTCACGTTTCGAAGACCCACCGAAGGA 253

Db 449 AAAGGAAGCGCAATAGACACCTGTTCAAGTCATCAGTCACGTTTCGAAGACCCACCGAAGGA 508

Qy 254 AAAGATCCAGAGATATAGAGATGATGAGGAGGGTCACTGATCTGTCAAAAGTGAGAGAG 313

Db 509 AAAGATCCAGAGATATAGAGATGATGAGGAGGGTCACTGATCTGTCAAAAGTGAGAGAG 568

Qy 314 TTCTAAGACGAAGATATGAAATCGTGGACACTTTGGGTGAAGAGCGCTTTGGCAAAGTTG 373

Db 569 TTCTAAGACGAAGATATGAAATCGTGGACACTTTGGGTGAAGAGCGCTTTGGCAAAGTTG 628

Qy 374 TAGAGTGCATTGATCATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGGCAT 433

Db 629 TAGAGTGCATTGATCATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGGCAT 688

Qy 434 TAGGCCGTTTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTA 493

Db 689 TAGGCCGTTTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTA 748

Qy 494 CTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGTGTCCAGATGTGTTCGATCATCATGTC 553

Db 749 CTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGTGTTCGATCATCATGTCGTC 808

Qy 554 ATGTTTGTATTGTTGTAACCTACTGGGACTTAGTACTTAGATTTCAATTAAGAAACA 613

Db 809 ATGTTTGTATTGTTGTAACCTACTGGGACTTAGTACTTAGATTTCAATTAAGAAACA 868

Qy 614 GCTTTTCGCCATTTCAAATTCACACATCAGCAGATGGCGTATCAGATCTGCCAGTCAA 673

Db 869 GCTTTTCGCCATTTCAAATTCACACATCAGCAGATGGCGTATCAGATCTGCCAGTCAA 928

Qy 674 TAAATTTTTTACATACATAATAAATTAACCCATACAGATCTGAAGCCTGAAAATATTTTGT 733

Db 929 TAAATTTTTTACATACATAATAAATTAACCCATACAGATCTGAAGCCTGAAAATATTTTGT 988

Qy 734 TTGTGAAGTCTGACTATGTAGTAAAATATAATTTCAAATGAAACGCTGATGACGACAC 793

Db 989 TTGTGAAGTCTGACTATGTAGTAAAATATAATTTCAAATGAAACGCTGATGACGACAC 1048

Qy 794 TGAAGAACACAGATATCAAAGTCTGTGACTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 853

Db 1049 TGAAGAACACAGATATCAAAGTCTGTGACTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1108

Qy 854 ACAGTACTTTTGGTGTCTACCCGGCAGCTACAGAGCTCCCGAGGTCATTTTGGCTTTAGGTT 913

Db 1109 ACAGTACTTTTGGTGTCTACCCGGCAGCTACAGAGCTCCCGAGGTCATTTTGGCTTTAGGTT 1168

Qy 914 GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCAATTTTGAATATTACCTTGGTT 973

Db 1169 GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCAATTTTGAATATTACCTTGGTT 1228

Qy 974 TCACAGTCTTTTCAGACTCATGATAGTAAGAGCAGCTGGCAATGATGGAACGAATATTAG 1033

Db 1229 TCACAGTCTTTTCAGACTCATGATAGTAAGAGCAGCTGGCAATGATGGAACGAATATTAG 1288

Qy 1034 GACCCATACCAACACACATGATTTCAGAAAAACAAGAAAAACGAAGTATTTCACCATAAACC 1093

Db 1289 GACCCATACCAACACACATGATTTCAGAAAAACAAGAAAAACGAAGTATTTCACCATAAACC 1348

Qy 1094 AGCTAGATTGGGATGAACACAGTTCTGCTGGTAGATATGTTAGGAGACGCTGCACAAACCGT 1153

Db 1349 AGCTAGATTGGGATGAACACAGTTCTGCTGGTAGATATGTTAGGAGACGCTGCACAAACCGT 1408

Qy 1154 TGAAGGAATTTATGCTTTGTCATGATGAGAAACATGAGAAATCTTTGACCTGGTTCGAA 1213

Db 1409 TGAAGGAATTTATGCTTTGTCATGATGAGAAACATGAGAAATCTTTGACCTGGTTCGAA 1468

Qy 1214 GAATGTTAGGAATATGATCCAACTCAAAGAAATTTACCTTGGATGAAGCATTTGACGATCCTT 1273

Db 1469 GAATGTTAGGAATATGATCCAACTCAAAGAAATTTACCTTGGATGAAGCATTTGACGATCCTT 1528

Qy 1274 TCCTTGTACTTATTAAAAAAGAAATGAAA 1301

Db 1529 TCTTTGACTTATTAATAAAAAAAAAAAAAA 1556

RESULT 3
AY408252 1446 bp DNA linear GSS 15-DEC-2003
LOCUS Homo sapiens CLK4 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY408252
VERSION AY408252.1 GI:39764223
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1446)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1446)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source location/Qualifiers
1..1446
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1446
/gene="CLK4"
/locus_tag="HCW3162"

ORIGIN
Query Match 52.9%; Score 1194; DB 9; Length 1446;
Best Local Similarity 92.8%; Pred No. 3.2e-246;
Matches 1194; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
Qy 14 GTCATTATTTAGAACAAAGGTCCTTGAATGACGAGATTATCGGACCGGAGATACGTTG 73
Db 161 GTCATTATTTAGAACAAAGGTCCTTGAATGACGAGATTATCGGACCGGAGATACGTTG 220
Qy 74 ACGAATACAGGAATGACTCTGTGAGGATATGTTCTTAGACATATTCACAGACATTTG 133
Db 221 ACGAATACAGGAATGACTCTGTGAGGATATGTTCTTAGACATATTCACAGACATTTG 280
Qy 134 AAAGCGGTATCGAATCCACTGCAGTAAATCTTCAGTCCGACGAGAGAAAGCACTCTTA 193
Db 281 AAAGCGGTATCGAATCCACTGCAGTAAATCTTCAGTCCGACGAGAGAAAGCACTCTTA 340
Qy 194 AAAGGAACGCAATAGACACTGTTCAAGTCACTAGTCACTGTTGGAAGACGCCACGAAGA 253
Db 341 AAAGGAACGCAATAGACACTGTTCAAGTCACTAGTCACTGTTGGAAGACGCCACGAAGA 400
Qy 254 AAAGATCCAGGATATAGAGATGATGAGGGGTCACTGATCTGTCAAAAGTGAGACG 313
Db 401 AAAGATCCAGGATATAGAGATGATGAGGGGTCACTGATCTGTCAAAAGTGAGACG 460
Qy 314 TTCTAAGAGCAAGATGAATGAAATCGTGGACACTTTTGGGTGAAGGAGCCCTTGGCAAAAGTTG 373
Db 461 TTCTAAGAGCAAGATGATNN 520
Qy 374 TAGAGTGCAATTGATCATGGCATGGATGGCATGATGTAGCAGTGAATAATCGTAAAAAATG 433

Db 521 NNNNNNNNNNNNNNNNNNNNNNGATGGCATGCTAGCAGTGAAAATCGTAAAAAATG 580
Qy 434 TAGGCCGTTTACCGTGAAGCAGCTCGTTTCAAGAAATCCAAAGTATTAGAGCACCTTAAATAGTA 493
Db 581 TAGGCCGTTTACCGTGAAGCAGCTCGTTTCAAGAAATCCAAAGTATTAGAGCACCTTAAATAGTA 640
Qy 494 CTGATCCCAATAGTGTCTTCCGATGTGTCAGATGCTAGAAATGTTTGTGATCATCATGTC 553
Db 641 CTGATCCCAATAGTGTCTTCCGATGTGTCAGATGCTAGAAATGTTTGTGATCATCATGTC 700
Qy 554 ATGTTTGTATGTTGTTTGAACCTAGTGGGACTTAGTACTTACGATTTTCAATTAAGAAAAACA 613
Db 701 ATGTTTGTATGTTGTTTGAACCTAGTGGGACTTAGTACTTACGATTTTCAATTAAGAAAAACA 760
Qy 614 GCTTTCTGCCATTTTCAAAATGACACATCAGGACAGATGGCGTATCAGATCTGCCAGTCAA 673
Db 761 GCTTTCTGCCATTTTCAAAATGACACATCAGGACAGATGGCGTATGGCGNNNNNNNNNNNNNNNN 820
Qy 674 TAAATTTTTTACATCATAATAAATTAACCCATACAGATCTGAAGCCTGAAATATTTTGT 733
Db 821 NNNNNNTTTTACATCATAATAAATTAACCCATACAGATCTGAAGCCTGAAATATTTTGT 880
Qy 734 TTGTGAAGTCTGACTATGTAGTCAAAATATATTTCTAAATGAAACGATGATGAAGCAGAC 793
Db 881 TTGTGAAGTCTGACTATGTAGTCAAAATATATTTCTAAATGAAACGATGATGAAGCAGAC 940
Qy 794 TGAATAAACACAGATATCAAAGTGTGTCATCTTGAAGTGCACCTGATGATGATGAATCAATC 853
Db 941 TGAATAAACACAGATATCAAAGTGTGTCATCTTGAAGTGCACCTGATGATGATGAATCAATC 1000
Qy 854 ACAGTACTTTGGTGTCTACCCGGCACTACAGAGTCCCGAGGTCAATTTGGCTTTAGGTT 913
Db 1001 ACAGTACTTTGGTGTCTACCCGGCACTACAGAGTCCCGAGGTCAATTTGGCTTTAGGTT 1060
Qy 914 GGTCTCAGCCTTGTGATGTTTGGAGCATAGTGTGCAATCTTTTGAATATTTACCTTGGTT 973
Db 1061 GGTCTCAGCCTTGTGATGTTTGGAGCATAGTGTGCAATCTTTTGAATATTTACCTTGGTT 1120
Qy 974 TCACAGTCTTTTTCAGACTCATGATAGTAAAGACACCTGGCAATGATGGAACGAATATTAG 1033
Db 1121 TCACAGTCTTTTTCAGACTCATGATAGTAAAGACACCTGGCAATGATGGAACGAATATTAG 1180
Qy 1034 GACCCATACCAACACATGATTCAGAAAAACAGAAAAACGCAAGTATTTTCAACATAACC 1093
Db 1181 GACCCATACCAACACATGATTCAGAAAAACAGAAAAACGCAAGTATTTTCAACATAACC 1240
Qy 1094 AGCTAGATTGGGATGAACACAGTTCCTGCTGTAGATATGTTAGGAGCGCTGCAACCGT 1153
Db 1241 AGCTAGATTGGGATGAACACAGTTCCTGCTGTAGATATGTTAGGAGCGCTGCAACCGT 1300
Qy 1154 TGAAGGAATTTATGCTTTGTCTCATGATGAAGAAACATGAAAACTGTTTCACTGGTTCGAA 1213
Db 1301 TGAAGGAATTTATGCTTTGTCTCATGATGAAGAAACATGAAAACTGTTTCACTGGTTCGAA 1360
Qy 1214 GAATTTAGAAATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCAATTCAGCATCCTT 1273
Db 1361 GAATTTAGAAATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCAATTCAGCATCCTT 1420
Qy 1274 TCTTTGACTTATTAATAAAGAAATGA 1299
Db 1421 TCTTTGACTTATTAATAAAGAAATGA 1446

RESULT 4
AF212224 1865 bp mRNA linear HTC 22-MAY-2001
LOCUS Homo sapiens CLK4 mRNA, complete cds.
DEFINITION AF212224
ACCESSION AF212224.1 GI:9437514
VERSION AF212224.1
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 1865)
Zhao, M., Song, H., Li, N., Peng, Y., Han, Z. and Chen, Z.

AUTHORS

A novel gene expressed in human bone marrow

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 1865)
Zhao, M., Song, H., Li, N., Peng, Y., Han, Z. and Chen, Z.

AUTHORS

Direct Submission

JOURNAL

Submitted (08-DEC-1999) Chinese National Human Genome Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, People's Republic of China
Location/Qualifiers

FEATURES

source

1..1865
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="bone marrow"
154..1515
/codon_start=1
/product="CLK4"
/protein_id="AAF87326.1"
/db_xref="GI:9437515"
/translation="MRHSKRTYCPDWDDKWDYKWRSSSHKRRKSRSSSAQENKRC
KINSMCDHYLERSINEDYHRRYIDSTRNDYTGCEPHRRQDRHESYQNHSS
KSSRGSRSSYKSHRIHSHSRSSHDEIVDTLGEAFGKVECDHKAGGRHVAV
KIVNDRYCAAAARSEIOVLEHLANTDPNSFRVCVQMLEWFEHGHICIVFELLGLST
YDFIKENGLFPLRDLHKMAVOICKSVNFLSHNKLTHDLKPENILPVQDPLEYAVN
PKIKEDERTLNPDIKVVDVFGSATYDDSHSLTVSTRHYRAPEVILALGNQPDGVNS
IGGILNEYLFQVTHDSKHLAMERILGPIPOHMIQKTRKRYKPHNQLDWEH
SSAGRYVRRRCRKLPEFMLCHDEHEKFLDLVRMLLEYDPTQRTLDALQHPFFDL
KKK"

CDS

49..61; Score 1118.8; DB 3; Length 1865;
Best Local Similarity 90.2%; Pred. No. 5.1e-230; Indels 0; Gaps 0;
Matches 1191; Conservative 0; Mismatches 129; Indels 0; Gaps 0;
329 ATGAATCGTGACACTTTGGTGAAGGAGCGCTTTGGCAAAAGTTCTAGAGTGCAATTGATC 388
|||||
545 ATGAATTTGTTGATCTTTAGTGAAGGAGCTTTTGGAAAGTTGTGGAGTGCAATGATC 604
|||||
389 ATGGATGGATGGCATGTAGCATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 448
|||||
605 ATAAAGCGGAGGTAGACATGTAGCAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 664
|||||
449 AAGCAGCTCGTTCAGAAATCCAAGTATTAGACACTTAAAGTACTGATCCCAATAGTG 508
|||||
665 AAGCTGCTCGCTCAGAAATACAAGTCTCGAAACATCTGAATACACAGACCCCAACAGTA 724
|||||
509 TCTTCGATGTGTCAGATGCTAGAAATGTTTGAATCATCATGCTCATGTTGTTGTTGTTGTT 568
|||||
725 CTTTCGCTGTGTCAGATGTTGAAATGTTGAAATGTTGAAATGTTGAAATGTTGAAATGTTG 784
|||||
569 TTGAACACTGGGACTTAGTACTTACGATTTTCATTAAGAAACACAGCTTTCTGCCATTTC 628
|||||
785 TTGAACACTGGGACTTAGTACTTACGATTTTCATTAAGAAACACAGCTTTCTGCCATTTC 844
|||||
629 AAATTGACCATCATGAGGAGATGGGATTCAGATCTGCCAGTCAATTAATTTTACATC 688
|||||
845 GACTGGATCATATCAGAAAGATGGCATATCAGATATGCAAGTCTGTGAATTTTTCGACA 904
|||||
689 ATAAATAATTAACCATACAGATCTGAAGCTGAAATATTTGTTGTTGTTGTTGTTGTTGTTG 748
|||||
905 GTAATAAGTTGACTCACAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAG 964
|||||
749 ATGTAGTCAAAATATATTTCTAAATGAACAGTGAACGACATCTGAAGAAACACAGATA 808
|||||
965 ACACAGAGCGGTATATCCCAATTAAGCGTGAACGACCTTATTAATCCAGATA 1024
|||||
809 TCAAGATTGTTGACTTTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 868
|||||
1025 TTAAGTTGTAGACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1084
|||||

ORIGIN

Query Match 49..61; Score 1118.8; DB 3; Length 1865;
Best Local Similarity 90.2%; Pred. No. 5.1e-230; Indels 0; Gaps 0;
Matches 1191; Conservative 0; Mismatches 129; Indels 0; Gaps 0;
329 ATGAATCGTGACACTTTGGTGAAGGAGCGCTTTGGCAAAAGTTCTAGAGTGCAATTGATC 388
|||||
545 ATGAATTTGTTGATCTTTAGTGAAGGAGCTTTTGGAAAGTTGTGGAGTGCAATGATC 604
|||||
389 ATGGATGGATGGCATGTAGCATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 448
|||||
605 ATAAAGCGGAGGTAGACATGTAGCAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 664
|||||
449 AAGCAGCTCGTTCAGAAATCCAAGTATTAGACACTTAAAGTACTGATCCCAATAGTG 508
|||||
665 AAGCTGCTCGCTCAGAAATACAAGTCTCGAAACATCTGAATACACAGACCCCAACAGTA 724
|||||
509 TCTTCGATGTGTCAGATGCTAGAAATGTTTGAATCATCATGCTCATGTTGTTGTTGTTGTT 568
|||||
725 CTTTCGCTGTGTCAGATGTTGAAATGTTGAAATGTTGAAATGTTGAAATGTTGAAATGTTG 784
|||||
569 TTGAACACTGGGACTTAGTACTTACGATTTTCATTAAGAAACACAGCTTTCTGCCATTTC 628
|||||
785 TTGAACACTGGGACTTAGTACTTACGATTTTCATTAAGAAACACAGCTTTCTGCCATTTC 844
|||||
629 AAATTGACCATCATGAGGAGATGGGATTCAGATCTGCCAGTCAATTAATTTTACATC 688
|||||
845 GACTGGATCATATCAGAAAGATGGCATATCAGATATGCAAGTCTGTGAATTTTTCGACA 904
|||||
689 ATAAATAATTAACCATACAGATCTGAAGCTGAAATATTTGTTGTTGTTGTTGTTGTTGTTG 748
|||||
905 GTAATAAGTTGACTCACAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAG 964
|||||
749 ATGTAGTCAAAATATATTTCTAAATGAACAGTGAACGACATCTGAAGAAACACAGATA 808
|||||
965 ACACAGAGCGGTATATCCCAATTAAGCGTGAACGACCTTATTAATCCAGATA 1024
|||||
809 TCAAGATTGTTGACTTTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 868
|||||
1025 TTAAGTTGTAGACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1084
|||||

QY 869 CTACCCGCGACTACAGAGCTCCCGAGGTCATTTTGGCTTTAGGTTGGTCTCAGCCTTGTG 928
|||||
D 1085 CTACAAGACATATTAGAGCAGCTCGAAGTATTTTAGCCCTAGGGTGGTCCAGCCTTGTG 1144
|||||
QY 929 ATGTTTGGAGCATAGTTGGCATCTTATTTGAATATTTACCTTGGTTTTCAGCTTTCAGA 988
|||||
D 1145 ATGTTTGGAGCATAGTTGGCATCTTATTAATGAATATTTACCTTGGTTTTCAGCTTTCAGA 1204
|||||
QY 989 CTCATGATAGTAAAGAGCAGCTGGCAATGATGGAACGAATATTAGGACCCATACACACAC 1048
|||||
D 1205 CTCATGATAGTAAAGAGCAGCTGGCAATGATGGAACGAATATTAGGACCCATACACACAC 1264
|||||
QY 1049 ACATGATTTCAGAAACAAAGAAACGCAAGTATTTTTCACCAATAACAGCTAGATTGGGATG 1108
|||||
D 1265 ACATGATTTCAGAAACAAAGAAACGCAAGTATTTTTCACCAATAACAGCTAGATTGGGATG 1324
|||||
QY 1109 AACACAGTTCTGCTGGTGTAGATATGTTAGGAGCGCTGCAACCCGTTGAAGGAATTTATGC 1168
|||||
D 1325 AACACAGTTCTGCTGGTGTAGATATGTTAGGAGCGCTGCAACCCGTTGAAGGAATTTATGC 1384
|||||
QY 1169 TTTGTGATGATGAAGAACATGAGAAACCTGTTTACCTGTTTGAAGAAATGTTAGAAATATG 1228
|||||
D 1385 TTTGTGATGATGAAGAACATGAGAAACCTGTTTACCTGTTTGAAGAAATGTTAGAAATATG 1444
|||||
QY 1229 ATCCAACTCAAGAAATTAACCTTGGATGAAGCAATTCGAGCACTCTTTTCTTGAATTTAA 1288
|||||
D 1445 ATCCAACTCAAGAAATTAACCTTGGATGAAGCAATTCGAGCACTCTTTTCTTGAATTTAA 1504
|||||
QY 1289 AAAAGAAATGAATGGGAATCAGTGGCTTCTATATATCTCTCTAGAAAGATTTACTTA 1348
|||||
D 1505 AAAAGAAATGAATGGGAATCAGTGGCTTCTATATATCTCTCTAGAAAGATTTACTTA 1564
|||||
QY 1349 AGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTTAAATTTTGTACA 1408
|||||
D 1555 AGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTTAAATTTTGTACA 1624
|||||
QY 1409 GTTAAGTGTAAATATTTGATGTTTGTATCAATAGCAATTAATTTCTTGAAGCAAGTAT 1468
|||||
D 1625 GTTAAGTGTAAATATTTGATGTTTGTATCAATAGCAATTAATTTCTTGAAGCAAGTAT 1684
|||||
QY 1469 GGTCTGTGATATGATTAAGAAATTAATAATTTTCTTTTGAATTTACCAATTTT 1528
|||||
D 1685 GGTCTGTGATATGATTAAGAAATTAATAATTTTCTTTTGAATTTACCAATTTT 1744
|||||
QY 1529 AAATACCTTTCAATATCTTTGTCAGGATTAATATGATGATCTTGCCTTTTGA 1588
|||||
D 1745 AAATACCTTTCAATATCTTTGTCAGGATTAATATGATGATCTTGCCTTTTGA 1804
|||||
QY 1589 CATGGAGTCACTCTGAAATGATTTTGTGAGTAAAGGAAATCTTGAATCTTTTATA 1648
|||||
D 1805 CATGGAGTCACTCTGAAATGATTTTGTGAGTAAAGGAAATCTTGAATCTTTTATA 1864
|||||

RESULT 5

AK013974

LOCUS

DEFINITION

AK013974

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AK013974 1629 bp mRNA linear HTC 03-APR-2004
Mus musculus 13 days embryo head cDNA, RIKEN full-length enriched
library, clone:3110002D09 product:cdc like kinase 4, full insert
sequence.
AK013974
AK013974.2 GI:26379557
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

QY 733 TTGTGAAGTCTGACTATGATCAATATATTTCTAAATGAACGTCATGAACGCACA 792
 Db |||||
 QY 713 TTGTGAAGTCTGACTATGATCAATATATTTCTAAATGAACGTCATGAACGCACA 772
 Db |||||
 QY 793 CTGAAGAAACACAGATATCAAGATTTGCTGCTTTGAAGTGAACGTCATGATGAACAT 852
 Db |||||
 QY 773 TTGAAGAAACACAGATATCAAGATTTGCTGCTTTGAAGTGAACGTCATGATGAACAT 832
 Db |||||
 QY 853 CACAGTACTTTTGTCTCTACCCGCGCACTACAGAGCTCCGAGGTCAATTTTGGCTTTAGT 912
 Db |||||
 QY 833 CATAGTACTTTTGTCTCTACCAAGGCACTACAAGGCTCCAGAGTCAATTTTGGCTTTAGT 892
 Db |||||
 QY 913 TGGTCTCAGCCTTGTGATTTTGGAGCATAGTGTGCATCTCTTATTAATATATACCTTGGT 972
 Db |||||
 QY 893 TGGTCTCAGCCTTGTGATTTTGGAGCATAGTGTGCATCTCTTATTAATATATACCTTGGG 952
 Db |||||
 QY 973 TTCACAGTCTTTTACAGCATCATGATAGTAAGAGCACTTGGCAATGATGGACGATATTA 1032
 Db |||||
 QY 953 TTCACAGTCTTTTACAGCACCACATAGTAAGAGCACTTGGCAATGATGGAGCGGATCTTA 1012
 Db |||||
 QY 1033 GGACCCATACCAACACATGATTCAGAAACCAAGAAACGCAAGTATTTTCAACATTAAC 1092
 Db |||||
 QY 1013 GGACCCATCCAGCAGCATATGATCCAGAGACCAAGAAACGCAAGTATTTTCAACATTAAC 1072
 Db |||||
 QY 1093 CAGCTAGATTTGGATGAACACAGTTCTCTCTGTAGATATGTTTGGAGACGCTGCAAAACCG 1152
 Db |||||
 QY 1073 CAGCTAGATTTGGAGCAGCATATGTTTGGAGCAGATGTTTGGAGACGCTGCAAAACCG 1132
 Db |||||
 QY 1153 TTGAAGGAATTTATCTTTGTCATGATGATGAAGCAATGAGAAACCTGTTGACCTGTTGGA 1212
 Db |||||
 QY 1133 TTAAAGGAATTTATCTTTGTCATGATGATGAAGCAATGAGAAACCTGTTGACCTGTTGGA 1192
 Db |||||
 QY 1213 AGAATGTTAGATATGATCCCACTCAAGAAATTAACCTTGGATGAAGCAATGAGCAATCCT 1272
 Db |||||
 QY 1193 AGAATGTTGGATATGATCCAGCAGAGAGATCACTTGGATGAAGCAATGAGCAATCCT 1252
 Db |||||
 QY 1273 TTCTTTGACTTATTAAGAAAGAAATGAG-TGGAGTCAAGGCTCTT-CTGTGACTTCTC 1332
 Db |||||
 QY 1333 TAGAAGATTAACCTT-AGACGTGTGTGATGATCAATTAACATCTTAATATTTTGTGAACA 1391
 Db |||||
 QY 1311 TAGGAGCACTTACTTCCAGACTGTGTGATCACTCAATTAACCTTCAATATTTTGTGAACA 1370
 Db |||||
 QY 1392 TTAATATTTTGTACAGTTAAGTCAATATTTGATGTTTGTGATCAATAGCAATTA 1451
 Db |||||
 QY 1371 TTAATATTTTGTACAGTTAAGTCAATATTTGATGTTTGTGATCAATAGCAATTA 1430
 Db |||||
 QY 1452 ACTTGTTAAGCAAGTATGCTCTTGATATGCAATTAAGAAATTAATATTTTCTTTT 1511
 Db |||||
 QY 1431 CTTTGTTAAGTATGCTGTGATATGAATGAATGAATGAATGAATGAATTTTCTTTT 1490
 Db |||||
 QY 1512 T--TGAATTAACATTTTAAATACCTTTGAAATATCCCTTTGTGTCAGTGAATATGT 1569
 Db |||||
 QY 1491 TGAATTTAATTTGCCACTTTAAGGCTTTTAGAATGCCCTTTTGTGTCAGTGAATATGT 1550
 Db |||||
 QY 1570 ATTGATCTTGTCTTTGATGATGAGGTCACCTCTGAAAGTCAATTTTCTT-----GAG 1622
 Db |||||
 QY 1551 ATTGTCCTCACTTTTGTGATGAAGGTTGACTCTGAAGTGAATTTTCTTCTCAAG 1610
 Db |||||
 QY 1623 TAAAGGAAATCTTGAC 1639
 Db |||||
 QY 1611 TAAAGGAAATCTTGAC 1627
 Db |||||

RESULT 6
 AY408254
 LOCUS AY408254 1421 bp DNA linear GSS 15-DEC-2003
 DEFINITION Mus musculus CLK4 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCESSION AY408254
 VERSION AY408254.1 GI:39764225

KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 1421)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source
 1..1421
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 <1..>1421
 /gene="CLK4"
 /locus_tag="HCM3162"
 gene
 ORIGIN
 Query Match 43.4%; Score 978.2; DB 9; Length 1421;
 Best Local Similarity 84.1%; Pred. No. 9e+200;
 Matches 1081; Conservative 0; Mismatches 180; Indels 25; Gaps 1;
 QY 14 GTCAATTATTTAGAAGCAAGTCTTGAATAGCGAGATATTCGGACCGGAGATACGTTG 73
 Db |||||
 QY 161 GTCAATTATTTAGAAGCAAGTCTTGAATAGCGAGATATTCGGACCGGAGATACATG 220
 Db |||||
 QY 74 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTATCACAGACATTTG 133
 Db |||||
 QY 221 ATGAATACAGAAATGACTACTTGCAGAGGATATGTTTCCAAGACATTTACCATGAGAG 280
 Db |||||
 QY 134 AAACGGGTATCGAATCCATCGAGTAAATCTTCAGTCCGAGCGAGGAGAGAGCATCTCTA 193
 Db |||||
 QY 281 AAAGCACTTACCGGATCCATTCGAGTAAATCTTCAGTCAGGAGCAGGAGAGCGCCCTA 340
 Db |||||
 QY 194 AAAGGAAGCGCAATAGACACTGTTCAAGTCATCAGTCAGTTCGAAAGAGCCACCGAAGCA 253
 Db |||||
 QY 341 AGAAGAAAGCTAATAGAACCTTGTGCAAGTCATCAGTCGATTCGAAAGAGCCACCGAAGCA 400
 Db |||||
 QY 254 AAAGATCCAGGATATAGAGGATGATGAGGAGGTTCACCTGATCTGTCAAAAGTGGAGACG 313
 Db |||||
 QY 401 AAAGATCCAGGATATAGAGGATGATGAGGAGGTTCACCTGATCTGTCAAAAGTGGAGACG 460
 Db |||||
 QY 314 TTCTAAGACGAAGATATGAATCTGGACACATTTGGGTGAAGGAGCCCTTTGGCAAAAGTTG 373
 Db |||||
 QY 461 TTCTAAGACGAAGATATGAATCTGGACACATTTGGGTGAAGGAGCCCTTTGGCAAAAGTTG 520
 Db |||||
 QY 374 TAGAGTCAATTGATCATGGCATGCGATGCGATGATGAGTGAAGTGAAGTGAAGTGAAGTGA 433
 Db |||||
 QY 521 NNN 580
 Db |||||
 QY 434 TAGGCCGTTTACCGTGAAGCAGCTCGTTTCAAGAAATCCAGTATTTAGAGCACTTTAAATAGTA 493
 Db |||||
 QY 581 TAGGCCGTTTACCGGAGGAGCAGCTCGTTTCAAGAAATCCAGTATTTAGAGCACTTTGAACGA 640
 Db |||||
 QY 494 CTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGCTAGAAATGTTTGTGATCATCATGTC 553
 Db |||||
 QY 641 CTGACCCCAACAGTGTCTTCCGATGCTGCCAGATGCTAGAGTGGTTTGTGATCATCATGTC 700
 Db |||||

QY	734	TTGTGAAGTCTGACTATGTAGTCAAAATATAATCTTAAATGAAACGCTGATGAACACAC	793
Db	881	NTNNNANGTCTGACTATGTNGTCAAAATATAATCTTAAATGAAACGCTGATGAACGNNN	940
QY	794	TGAAAAACACAGATATCAAAAGTGTGGACTTTGGAGTGGCAAGTGCACGATGATGATGAACATC	853
Db	941	TGAAAAACACAGATATCAAAAGTGTGGACTTTGGAGTGGCAAGTGCACGATGATGATGAACATC	1000
QY	854	ACAGTACTTTGGTGTCTACCCGGCACTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGGTT	913
Db	1001	ACAGTACTTTGGTGTCTACCCGGCACTACAGAGCTCCCGAGGTCAATTTTGGCTTTAGGTT	1060
QY	914	GGTCTCAACCTTTGGAGCATAGGTGTGCAATTTCTTATTTGAATATTAACCTTGGTT	973
Db	1061	GGTNNCNGCCTTGTGATGTTTGGAGCATAGGTGTGCAATTTCTTATTTGAATATTAACCTTGGTT	1120
QY	974	TCACAGTCTTTTCAGACTCATGATAGTAAAGACACCTGCGCAATGATGAAAGATATTAG	1033
Db	1121	TCACAGTCTTTTCAGACTCATGATAGTAAAGACACCTGCGCAATGATGAAAGATATTAG	1180
QY	1034	GACCCATACCAACACACATGATTTTCAAGAAAACAGAAACGCAAGTATTTCACCAATACC	1093
Db	1181	GACCCATACCAACACACATGATTTTCAAGAAAACAGAAACGCAAGTATTTCACCAATACC	1240
QY	1094	AGCTAGATTGGATGAACACACAGTCTCTGCTGTGATATGTTAGGAGACGCTGCAACCGT	1153
Db	1241	NN	1300
QY	1154	TCAAGGAATTTATGCTTGTCTATGATGAAGACATGAGAACTGTTTCAACCTGGTTCGA	1213
Db	1301	TCAAGGAATTTATGCTTGTCTATGATGAAGACATGAGAACTGTTTCAACCTGGTTCGA	1360
QY	1214	GAAATGTAAGATATGATCAACTCAAGAAATTAACCTTGGATGAAGCAATTCAGCATCCTT	1273
Db	1361	NN	1420
QY	1274	TCTTTGACTTATTAAGAAAGAAATCA	1299
Db	1421	TCTTTGACTTATTAAGAAAGAAATCA	1446
RESULT 8			
CR591789			
LOCUS	full-length cDNA clone	CL0BA009ZE09	linear HTC 21-JUL-2004
DEFINITION			
ACCESSION	CR591789		
VERSION	CR591789.1	GI:50472596	
KEYWORDS	HTC; CNSLT cDNA		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Li, W. B., Gruber, C., Jessee, J. and Polayes, D.		
TITLE	Full-length cDNA libraries and normalization		
JOURNAL	Unpublished		
REMARK	Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/		
REFERENCE			
AUTHORS	Genoscope.		
TITLE	Genoscope.		
JOURNAL	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)		
COMMENT	1st strand cDNA was primed with a NotI-Oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.		
FEATURES	Location/Qualifiers		

source	1. .1737	/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="CL0BA009ZE09"	
		/tissue_type="Placenta"	
		/plasmid="pCMVSPORT_6"	
ORIGIN			
Query Match	36.1%	Score 814.6; DB 3; Length 1737;	
Best Local Similarity	73.4%	Pred. No. 1.4e-164;	
Matches 1102; Conservative	0; Mismatches 382; Indels 18; Gaps 4;		
QY	13	AGTCATTTATTTAGAAAGCAAGTCTTGAATGAGCGAGATTATTCGGGACCGGAGATACGTT	72
Db	238	AGCCATTTATTTGGAAAGCAGGCTCTATAAATGAGAAAGATTATCATAGTCGACGCTACATT	297
QY	73	GACCAATACAGGAATGACTACTGTGAGGATATGTTCTCTAGACATTATTCACAGAGACATT	132
Db	298	GATGAGTACAGAAATGACTACACCAAGGATGTGAACCTGGACATCGCCAAAGAGACCAT	357
QY	133	GAAGCGGTATCGAATCCACTGCGAGTAAATCTTCAGTCCGAGCAGGAGAGCAGTCT	192
Db	358	GAAGCGGTATCAGAACCATAGTAGCAAGTCTTCTGCTAGAGTGGAGAGTAGTTAT	417
QY	193	AAAAAGGAGCGCAATAGACAC-----TGTTCAAAGTCACTAGTCACTGTTTCAAGAGCCAC	246
Db	418	AAAAAGCAACACAGGATTCACACAGACTCTTCACATCGTCGTTCACTGGGAAGAGTCCAC	477
QY	247	CGAAGGAAAGATCCAGGATATAGAGGATGATGAGGAGGCTCACCTGATCTGTCAAAAT	306
Db	478	CGAAGGAAAGAACCCAGGAGTGTAGAGGATGATGAGGAGGCTCACCTGATCTGTCAAAAT	537
QY	307	GGAGAGCTTCTAAGAGCAAGATATGAAATCGTGGACACTTTGGGTGAAGAGGAGCTTTGGC	366
Db	538	GGAGAGCTTCTAAGAGCAAGATATGAAATCGTGGACACTTTGGGTGAAGAGGAGCTTTGGC	597
QY	367	AAAGTTGTAGAGTGCAATGATCATGGCATGGCATGGCATGGCATGGCATGGCATGGATCGTA	426
Db	598	AAAGTTGTAGAGTGCAATGATCATGGCATGGCATGGCATGGCATGGCATGGATCGTA	657
QY	427	AAAAATGTAGCGCTTACCGTGAAGCAGCTCGTTTCAGAAATCCAGATATTAGAGCATT	486
Db	658	AAAAATGTAGCGCTTACCGTGAAGCAGCTCGTTTCAGAAATCCAGATATTAGAGCATT	717
QY	487	AATAGTACTGATCCCAATAGTGTCTTCGATGTGTCCAGATGCTAGATGTTGTTGATCAT	546
Db	718	AATACACAGAGCCCAACAGTACTTTCCGCTGTGTCCAGATGTTGGATGGTTTGAGCAT	777
QY	547	CATGTCATGTTTGTATTTGTGTTTGAACCTACTGGGACTTTAGTACTTACGATTTCAATAA	606
Db	778	CATGTCATGTTTGTATTTGTGTTTGAACCTACTGGGACTTTAGTACTTACGATTTCAATAA	837
QY	607	GAAGACAGCTTTCTGCCATTTCAAAATGACACATCAGGAGATGCGGCTATCAGATCTGC	666
Db	838	GAAGACAGCTTTCTGCCATTTCAAAATGACACATCAGGAGATGCGGCTATCAGATCTGC	897
QY	667	CAGTCAATAAATTTTTCATCATATAAATTAACCCATACAGATCTCAAGGCTCAAAAT	726
Db	898	AAGTCTGTAATTTTTCACAGTAATTAAGTTGACTCAACAGACTTAAGGCTCAAAAT	957
QY	727	ATTTTGTGTAAGTCTGACTATGTTAGTCAAAATATAAATTTCAAAATGAAGCTGATGAA	786
Db	958	ATTTTGTGTAAGTCTGACTATGTTAGTCAAAATATAAATTTCAAAATGAAGCTGATGAA	1017
QY	787	CGACACTGAAAAACACAGATATCAAGTGTGTTGCTTTGGAAGTGCACATGATGAT	846
Db	1018	CGACACTGAAAAACACAGATATCAAGTGTGTTGCTTTGGAAGTGCACATGATGAT	1077
QY	847	GAAATACAGTACTTTGCTGTACCCGGCACTACAGAGCTCCCGAGGCTCATTTTGGCT	906
Db	1078	GAAATACAGTACTTTGCTGTACCCGGCACTACAGAGCTCCCGAGGCTCATTTTGGCT	1137

```
QY 907 TTAGTGTGCTCAGCCTTGATGTTTGGAGCAGTAGGTGCAATCTTTATTGAATATTAC 966
Db 1138 CTAGGGTGGTCCCAACCATGATGCTCTGGAGCAGTAGGATGCAATCTTTATTGAATATTAC 1197
QY 967 CTTGGTTTTCAGAGCTTTTCAGACTCATGATAGTAAGAGCAGCTGGCAATGATGGAACGA 1026
Db 1198 CTTGGGTTTACCGTATTCTTCCAAACACACATAGTAGAGGAGCATTTAGCAATGATGGAAGG 1257
QY 1027 ATATTAGAGCACCATACCAACACATGATTCAGAAAACCAAGAAACGCAAGTATTTCAC 1086
Db 1258 ATCTTGGACCTCTACCAAAACATATGATACAGAAAACCAAGAAACGTAATAATTTTCAC 1317
QY 1087 CATAACACAGTAGATTGGATGCAACACAGCTTCTGCTGTAGATATGTTAGGAGAGCTGTC 1146
Db 1318 CACGATCGATTAGACTGGATGACACAGTCTGCGCGCAGATATGTTTCAAGAGCGCTGT 1377
QY 1147 AACCGTTGAAGGAATTTATGCTTTGTGTCATGATGAAGAAACATGAGAAATCTGTTGACCTG 1206
Db 1378 AACCTCTGAAGGAATTTATGCTTTCTCAAGATGTTGAACATGAGCGCTCTCTTTGACCTC 1437
QY 1207 GTTCGAAGAAATGTTAGATATGATCCAACTCAAAGAAATTTACTTGGATGAAGCAATTCGAG 1266
Db 1438 ATTCAAGAAATGTTGGAGTATGATCCAGCCAAAGAAATTTACTCTCAGAGAAAGCTTTAAG 1497
QY 1267 CATCCTTTCTTTGACTTATTAAGAAAGAAATGAAATGGGAATCAGTGGTCTTTACTATATA 1326
Db 1498 CATCCTTTCTTTGACTTCTGAAGAAA-----GTATATAGATCTGTAATTTGGACA 1548
QY 1327 CTTCTCAGAGAGATTTACTTAAGACTGTGTCAGTTC--AACTAAACATCTTAATAATTTTT 1384
Db 1549 GCTCTCTCGAAGAGA-TCTTACAGACTGTATCAGTCTAAATTTTAAATTTTAAAGTTATTT 1607
QY 1385 GTAACATTAATATTTTGTACAGTTAAGTAAATATTTGATGTTTGTGATCAATAGC 1444
Db 1608 TGTACAGCTTTGTAAATTTCTTAAATTTTATATTTGCAATGTTTATTTTGTGGTAAT 1667
QY 1445 ATAATTAATCTTTGTAAGCAAGTATGCTCTTGATATGCAATGATAGAAAATTAATAATTT 1504
Db 1668 TTGGTTTCAATTAAGTACATAGCTAAGGTAATGAACATCTTTTTCAGTAATTTGTAAGTGT 1727
QY 1505 TT 1506
Db 1728 TT 1729

RESULT 9
CR612115
LOCUS
DEFINITION
CR612115 1746 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CSODF036YH24 of Fetal brain of Homo sapiens
(human).
ACCESSION
CR612115.1 GI:50492922
VERSION
HTC; CNSL7_cDNA.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1746)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Paradise Avenue
2 (bases 1 to 1746)
Genoscope.
Direct Submission
TITLE
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
JOURNAL
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
```

```
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
1..1746
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODF036YH24"
/tissue type="Fetal brain"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 36.1%; Score 814.8; DB 3; Length 1746;
Best Local Similarity 73.4%; Pred. No. 1.4e-164;
Matches 1102; Conservative 0; Mismatches 382; Indels 18; Gaps 4;
QY 13 AGTCATTTATTAGAACGAGGTCTTGAATCAGCGAGATTTATCGGACCGAGATACGTT 72
Db 237 AGCCATTTATTGGAAAGCAGGTCTATAAATGAGAAAGATTTATCATATGTCGACGCTACAT 296
QY 73 GACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTTATCAACAGACAT 132
Db 297 GATGAGTACAGAAATGACTACCTCAAGGATGTGAACCTGACATCGCCAAAGACCAT 356
QY 133 GAAAGCGGTTATCGAATCCACTGTCAGTAAATCTTCAGTCCGACGAGAGAGAGAGTCCT 192
Db 357 GAAAGCGGTTATCAGAACCATAGTAGCAAGTCTTCTGTTGAAGTGGAAAGTAGTTAT 416
QY 193 AAAAGGAGCCCATAGACAC-----TGTTCAGTCTCATGTCAGTCCAGAGGCCAC 246
Db 417 AAAAGCAACACAGGATTCACACAGTACTTCACATCGTCTTACATGGGAAGAGTCC 476
QY 247 CGAAGGAAAGATCAGGAGTATAGAGGATGATGAGAGAGGCTCACCTGATCTGCAAAAGT 306
Db 477 CGAAGGAAAGAACACAGGAGTGTAGAGATGATGAGAGGCTCACCTGATCTGCAAGT 536
QY 307 GGAGACGTTCTTAAGACGAAGATATGAAATCTGTGGACACTTTGGGTGAAGAGCCCTTGGC 366
Db 537 GGAGACGTTACTAAGTGCAGATATGAAATCTGTGATCTTTAGGTGAAGAGGCTTTTGA 596
QY 367 AAGTGTGTAGAGTCATGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 426
Db 597 AAGTGTGTGAGTGTGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 656
QY 427 AAAAATGTAGCCGCTTACCGTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGACACTTA 486
Db 657 AAAAATGTGAGTATGATGCTGTGAGCTGCTCGCTCAGAAATACAAAGTCTGGAACATCTG 716
QY 487 AATAGTACTGATCCCAATAGTGTCTTCGGATGTGTCCAGATGCTAGAAATGGTTTGATCAT 546
Db 717 AATACAAACAGACCCCAACAGTACTTTCCGCTGTGTCCAGATGTTGGAAATGGTTTGAGCAT 776
QY 547 CATGTCATGTTGTATTGTGTTTGAATCTGAGCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAGCT 606
Db 777 CATGTCATGTTGTATTGTGTTTGAATCTGAGCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAGCT 836
QY 607 GAAACACAGCTTTCTGCCATTTTCAAAATGACCAATCAGCAGATGCGGTATCAGATCTGC 666
Db 837 GAAATGTTTCTTACCATTTTCCGATGATCATATCAGAAAGATGCGCATATCAGATATGCG 896
QY 667 CAGTCAATAAATTTTATCATCATTAATAATTAACCCATACAGATCTGCAAGCCCTGAAAT 726
Db 897 AAGTCTGTAATTTTGTGACAGTAAATGAGTGTGACTCACACAGACTTAAAGCCCTGAAAC 956
QY 727 ATTTGTTGTGAGAGTCTGACTATGCTCAAAATATATTTCTAAATCAAAACGATGAA 786
Db 957 ATCTTATTGTGAGCTGTGACTCACAGAGCGGTATTAATCCAAATATAAAGCGTATGNA 1016
QY 787 CGCACACTGAAACAAACACAGATATCAAAAGTGTGTGACTTTGGAAGTGCACACGATGAT 846
Db 1017 CGCACCTTAATAAATCCAGATATTAAGTGTGTGACTTTGTTAGTGTGCAACATATGATGAC 1076
QY 847 GAAACATCAGTACTTTGGTGTCTTACCCGCGCACTACAGAGCTCCGAGGTCAATTTGGCT 906
```

```

|||||
1077 GAACATCAGTACATGGTATCTACAAGACATTATAGACACCTGAAGTTATTATTGACC 1136
QY 907 TTAGGTGGTCTCAGCCTTGTGATCTTTGGAGCATAGGTGTGCATCTTTATTGAATATTAC 966
Db 1137 CTAGGGTGGTCCCAACCATGTGATCTCTGGAGCATAGGATGATCTTTATTGAATCTAT 1196
QY 967 CTGGTTTTCAGAGTCTTTTCAGATCATGATAGTAAAGAGCACCTGGCAATGATGAACGA 1026
Db 1197 CTGGGTTTACCGTATTTCCAAACACACGATAGTAGGAGCATTTAGCAATGATGAAAGG 1256
QY 1027 ATATTAGACCCATACCAACACATGATTCAGAAAACAGAAAACGCAAGTATTTTTCAC 1086
Db 1257 ATTCTTGGACCTCTACCAAAACATATGATACAGAAAACCAAGAAACGTAATATTTTTCAC 1316
QY 1087 CATAACACAGCTAGATGTTGGATGAACACAGATCTGCTGGTGTAGATATGTTAGGAGCGCTGC 1146
Db 1317 CACGATCATGATAGCTGGATGAACACAGATCTGCGGCGAGATGTTTTCAGAGCGCTGT 1376
QY 1147 AAACCGTTGAAGGAATTTATGCTTTGTCTATGATGAAGAACATGAGAAAACCTGTTGACCTG 1206
Db 1377 AAACCTCTGAAGGAATTTATGCTTTCTCAAGATGTTGAACATGAGCGTCTCTTTGACCTC 1436
QY 1207 GTTCGAAGATTTAGATATGATCCAACTCAAGAAATTCCTTGGATGAGCAATTCGAG 1266
Db 1437 ATTCAGAAAATGTTGGAGTATGATCCAGCCAAAGAAATTTACTCTCAGAGAGAGCCCTTAAAG 1496
QY 1267 CATCTCTTCTTGACTTATTAAGAAAGAAATGAAATGGAATCAGTGGTCTTACTATATA 1326
Db 1497 CATCTCTTCTTGACTTCTGAGAAA-----GTATATGATCTGTAATTTGGACA 1547
QY 1327 CTCTCTAGAGAGATTAATTAAGACTGTGTCAGTCT--AACTAAACATTTCTAATATTTTT 1384
Db 1548 GCTCTCTCGAGAGA-TCTTACAGACTGTATCAGTCTAATTTTAAATTTTAAAGTTATTT 1606
QY 1385 GTAACATTAATTTTGTACAGTTAAGTGAATTAATTTGTATGTTTGTATCAATATAGC 1444
Db 1607 TGTACAGCTTTGTAAATTTCTTAACATTTTATATTTGTCATGTTTATTTTGTGGTAA 1666
QY 1445 ATAATTAATCTTTGAAGCAAGTATGCTCTGATATGATGATGATGATGATGATGATGAT 1504
Db 1667 TTGGTTCAATTAAGTACATAGCTAAGTAATGAACATCTTTTTCAGTAATTTGAAGTAT 1726

QY 1505 TT 1506
Db 1727 TT 1728

RESULT 10
LOCUS CR617472
DEFINITION full-length cDNA clone CS0DI036YH19 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR617472
VERSION CR617472.1 GI:50498279
KEYWORDS HTC; CDSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1754)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 1754)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

```

```

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES             Location/Qualifiers
     source            1..1754
     organism="Homo sapiens"
     mol_type="mRNA"
     db_xref="taxon:9606"
     clone="CS0DI036YH19"
     tissue type="Placenta Cot 25-normalized"
     plasmid="pCMVSPORT_6"
ORIGIN
Query Match      36.1%; Score 814.8; DB 3; Length 1754;
Best Local Similarity 73.4%; Pred. No. 1.4e-164;
Matches 1102; Conservative 0; Mismatches 382; Indels 18; Gaps 4;
QY 13 AGTCATTATTTAGAAAGCAAGTCTTGAATGAGCGAGATTTATCGGAGCCGGAGATACGTT 72
Db 238 AGCCATTATTTGAAAAGCAGGTCTATAAATGAGAAAGATTATCATAGTCGACGCTACATT 297
QY 73 GACGATACAGAAATGACTACTGTGAGGATATGTTCTTACGACATTTATCAGAGACATT 132
Db 298 GATGAGTACAGAAATGACTACATCAAGGATGTGAACCTGGACATCGCCAAAGAGACCAT 357
QY 133 GAAAGCGGATTCGAATCCACTGTCAGTAAATCTTCAGTCCGAGCAGGAGAGAGAGTCTCT 192
Db 358 GAAAGCGGATTCAGAACCATAGTAGCAAGTCTTCTGGTAGAGTGGAAAGAGTAGTTAT 417
QY 193 AAAAGGAGCGCAATAGACAC-----TGTTCAAAGTCAATAGTCACTGTCGAGAGCCAC 246
Db 418 AAAAGCAACACAGATTTCACACAGACTTTTCAATCGTCTGTTTCAATGGAAGAGTCACT 477
QY 247 CGAAGAAAAGATCCAGAGTATAGAGATGATGAGGAGGCTCACCTGATCTGTCAAAGT 306
Db 478 CGAAGAAAAGAACCCAGAGTGTAGAGGATGATGAGGAGGCTCACCTGATCTGTGAGAGT 537
QY 307 GGAGAGCTTTCTAAGAGCAAGATATGAAATCGTGACACTTTGGGTGAAGAGGAGCTTTGGC 366
Db 538 GGAGAGCTTCTAAGTGAAGATATGAAATTTGTTGATCTTTAGGTGAAGAGGAGCTTTTGA 597
QY 367 AAAGTTGTAGAGTGCATTGATCATGGCATGGATGGATGTCATGTCATGTAGAGTGAATTCGTA 426
Db 598 AAAGTTGTGAGTGCATCGATCATAAAGCGGAGGTGACATGATGACAGTAAATAATAGTT 657
QY 427 AAAAATGTAGCGGTTTACCGTGAAGAGCTGTTTCAAGAAATCCAAAGTATTAGAGCACTTA 486
Db 658 AAAAATGTGATAGATAGTGTGAAAGCTGCTCGCTCAGAAATACAAAGTTCTGGAACATCTG 717
QY 487 AATAGTACTGATCCCAATAGTGTCTTCGATGTGTCAGATGTCAGATGTTGTTTTCATCAT 546
Db 718 AATCAACAGACCCCAACAGACTTTTCCGCTGTGTCAGATGTTGGATGTTTGGAT 777
QY 547 CATGTCATGTTTGTGTTTGTGTTTGAACACTACCTGGAGCTTTAGTACTTACGATTTCAATAA 606
Db 778 CATGTCATGTTTGTGTTTGTGTTTGAACACTTGGGACTTTAGTACTTACGATTTCAATAA 837
QY 607 GAAACAGCTTTTCTGCCATTTTCAAAATGACACATCAGGAGATGCGGCTATCAGATCTGC 666
Db 838 GAAAAATGTTTCTTACCATTTTCGACTGGATCATATCAGAAAGATGSCATATCAGATATGC 897
QY 667 CAGTCAATAAATTTTATACATCAATAAATAAATTAACCCATACAGATCTGAAGCTCAAAAT 726
Db 898 AAGTCTGTGAATTTTTCAGCAGTAATGAGTGTGACTCACAGACTTTAAGCCTGAAATAC 957
QY 727 ATTTTGTGTAAGTCTGACTATGTTAGTCAAAATATAATTTCTAAATGAACGTCATGAA 786
Db 958 ATCTTATTTGTGAGTCTGACTACACAGAGGCGTATATCCAAATATAAAGCTGATGAA 1017
QY 787 CGCACACTGAAACACAGATATCAAGTGTGTTGCTTTGAAAGTGAACGTCATGATGAT 846

```



```

Db      726  GATATGTTAGGAGCGCTGCAAGCGTTAAAGGAATTTATGCTGTGTATGACGAGGAGC 785
Qy      1187  ATGAGAACTGTTTGACCTGGTTTGAAGAATGTTAGAATATGATCAACTCAAGAAATTA 1246
Db      786  ATGAGAGCTGTTTGACCTGGTTC-AAAATGTTGGAGTAGACCCACCGAGAGATCA 844
Qy      1247  -CCTTGGATGAAGCAATGAGCATCCTTCTTTGACTATTATTAAGAAAGAAATGAATGGG 1305
Db      845  TCTTTGGATGAAGCAATGAGCATCCTTCTTCTGACTTTATTAAGAAAGAAATGAATGGG 904
Qy      1306  RATCATGCTTACTATATATCTCT-AGAGAGATTAAGATTAAGATTAAGATTAAGATTA 1364
Db      905  AGCCAGGGGGCTTCTGGGAAACTCCCTTAAGACAGGTTACTCCAGATGGGTGAGCCAA 964
Qy      1365  CTAACACTT 1373
Db      965  CTAACACTT 973

RESULT 12
ACCESSION AK076199
VERSION 1689 bp mRNA linear HTC 03-APR-2004
KEYWORDS Mus musculus 18 days pregnant adult female placenta and extra
SOURCE embryonic tissue cDNA, RIKEN full-length enriched library,
          (PROTEIN KINASE STY) homolog [Mus musculus], full insert sequence.
ORGANISM (PROTEIN KINASE STY) homolog [Mus musculus], full insert sequence.

REFERENCE 1
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
TITLE Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
TITLE Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
TITLE Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
JOURNAL Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
MEDLINE Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
PUBMED Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
        Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
        Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
        RIKEN integrated sequence analysis (RISA) system--384-format
        sequencing pipeline with 384 multipipillary sequencer
        Genome Res. 10 (11), 1757-1771 (2000)
        20530913
        11076861
        11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
TITLE PANTOM Consortium.
JOURNAL Functional annotation of a full-length mouse cDNA collection
MEDLINE Nature 409, 685-690 (2001)
PUBMED 11076861
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
TITLE Group Phase I & II Team.
JOURNAL Analysis of the mouse transcriptome based on functional annotation
MEDLINE of 60,770 full-length cDNAs
PUBMED Nature 420, 563-573 (2002)
REFERENCE 6
AUTHORS Nature 420, 563-573 (2002)

```

```

AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirose, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, K., Ohno, M., Ohnishi, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
FEATURES
source
    1..1689
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="C57BL/6J"
    /db_xref="PANTOM.DB:3830430K12"
    /db_xref="taxon:10090"
    /clone="3830430K12"
    /sex="female"
    /tissue_type="placenta and extra embryonic tissue"
    /clone_lib="RIKEN full-length enriched mouse cDNA library"
    /dev_stage="18 days pregnant adult"
    105..1556
    /note="PROTEIN KINASE CLK1 (EC 2.7.1.-) (CLK) (PROTEIN
    KINASE STY) homolog [Mus musculus] (SWISSPROT|P22518,
    evidence: FASTY, 95%ID, 100%length, match=1449)
    putative"
    1670..1675
    /note="putative"
    1689
    /note="putative"
    misc_feature
    105..1556
    /note="PROTEIN KINASE CLK1 (EC 2.7.1.-) (CLK) (PROTEIN
    KINASE STY) homolog [Mus musculus] (SWISSPROT|P22518,
    evidence: FASTY, 95%ID, 100%length, match=1449)
    putative"
    1670..1675
    /note="putative"
    polyA_signal
    1689
    /note="putative"
    polyA_site
    1689
    /note="putative"
    ORIGIN
    Query Match 32.5% Score 734.2; DB 3; Length 1689;
    Best Local Similarity 74.2%; Pred. No. 3.1e-147;
    Matches 957; Conservative 0; Mismatches 324; Indels 9; Gaps 2;
    Qy 12 CAGTCATTATTAGAAAGCAAGGTCCTTGAATGAGCGAGATATTCGGACCGGAGATACGT 71
    Db 263 CAGCTATTATCTGGAAGCAGATCCATAAATGAGAAAGCTTATCATAGTCGACGCTATGT 322
    Qy 72 TCAGCAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTTATCAGAGACAT 131
    Db 323 TCATGAATACAGGAATGACTACTATG- --GGCTACGAGCGAGGCGATCCCTATGGAGAAC 379
    Qy 132 TGAAGCGGGTATCGAATCCACTGCAGTAAATCTTCAGTCGCGACGAGAGAGAGTC 191
    Db 380 TGGAGCAGATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 439
    Qy 192 TAAAGGAAGCGCAATAG- ----ACACTGTTCAAGTCATCAGTCAGTCGAGAGGCA 245
    Db 440 ACMAAGTAACACAGGAGTCGCCACACACTTCGCGACGACCATTCACACGGAGAGTCA 499
    Qy 246 CCGAAGGAAGATCCAGGATGATGAGGATGATGAGGAGGTCACCTGATCTGTCAAG 305

```

Db 500 CCGAAGAAAATAAGCGAGGAGTGATAGAGATGATGAGGAGGTCACCTGATCTGTGACAG 559
 QY 306 TGGAGACGCTTCTAAGAGCAAGATATGAATCGTGACACTTTGGTGAAGGAGCCTTGG 365
 Db 560 TGGAGACGCTAAGTGAAGATATGATATGTTGATATCTTTAGGTGAAGGTCCTTGG 619
 QY 366 CAAAGTGTAGAGTGCATTTGATCATGGCATGATGGCATGATGATGATGATGATGATGAT 425
 Db 620 AGAAGTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 679
 QY 426 AAAAAATGAGCCGTTTACCGTGAAGAGAGCTGTTTCAAGAAATCCAAATGATGATGATGAT 485
 Db 680 TAAAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 739
 QY 486 AAATAGTACTGATCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 545
 Db 740 GTATACAGCAGACCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 799
 QY 546 TCATGCTCATGTTTGTATGTTGTTTGAATCTAGTGAATCTAGTGAATCTAGTGAATCTAGTGAAT 605
 Db 800 TCGAGATCAGATTTGATGTTGTTGATGTTTGAATCTAGTGAATCTAGTGAATCTAGTGAAT 859
 QY 606 AGAAACAGCTTTCGCAATTTCAAAATGACATCAGGAGATGAGGATGATGATGATGATGATGAT 665
 Db 860 GAAAAACAGTTTCTGCGGTCTCGAATGATCATATCAGGAGATGAGGATGATGATGATGATGATGAT 919
 QY 666 CAGTCAATTAATTTTACATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 725
 Db 920 CAATGCTGTAATCTTTTGCATAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 979
 QY 726 TATTTTGTGTTGAGAGTCTGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 785
 Db 980 CATCTTATTTGTGAGTCTGACTACAGAGGCTTATTAATCCAAATGAACGATGATGATGATGATGAT 1039
 QY 786 AGCACAATGAAAAACAAGATATCAAAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 845
 Db 1040 AGCTACTATGATTAATCAGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1099
 QY 846 TGAACATCAGACTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 905
 Db 1100 TGAACACACAGCAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1159
 QY 906 TTTAGTGTGTT 965
 Db 1160 CTTGCGGTGTT 1219
 QY 966 CTTGCGGTGTT 1025
 Db 1220 TCTTGGATTTACAGTTTTCGACTCATGATAGCAGGGAACATTTAGCAATGATGGAAG 1279
 QY 1026 AATATTAGACCCATACACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1085
 Db 1280 GATTTGTCACCATACCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1339
 QY 1086 CCATAACAGCTAGATTTGGATGAACACAGTTTCTGCTGTTGATGATGATGATGATGATGATGATGATGAT 1145
 Db 1340 TCATGATCGAT 1399
 QY 1146 CAAACCGTTGAAAGGAATTTATGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1205
 Db 1400 TAAACCTCTGAGGAGTTTATGCTATCTCAGGATGCGCAACATGATGATGATGATGATGATGATGATGAT 1459
 QY 1206 GGTTCGAGAAATGTTAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1265
 Db 1460 CATTGGGAAATGTTGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1519
 QY 1266 GCATCTCTTCTTTGACTTTATTTAAAAAGAA 1295
 Db 1520 GCATCTCTTCTTTTACCCTTTAAAAAGCA 1549

RESULT 13
 BM786882

LOCUS BM786882 722 bp mRNA linear EST 05-MAR-2002
 DEFINITION K-EST0065716 S10SNU1 Homo sapiens cDNA clone S10SNU1-17-D07 5',
 mRNA sequence.
 ACCESSION BM786882
 VERSION BM786882.1 GI:19135114
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (Bases 1 to 722)
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 TITLE 21C Frontier Korean EST Project 2001
 JOURNAL Unpublished (2002)
 COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 17 row: D column: 07
 High quality sequence stop: 722.
 FEATURES
 Location/Qualifiers
 1..722
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S10SNU1-17-D07"
 /sex="M"
 /tissue_type="Stomach"
 /cell_type="Lymphoblast-like"
 /cell_line="SNU-1"
 /lab_host="Top10F"
 /clone_lib="S10SNU1"
 /note="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI;
 Site 2: XhoI; The poly (A) + RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tobacco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including SfiI
 site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized with Superscript II using SfiI
 oligo-dT primer. After first strand synthesis, RNA was
 degraded by NaOH treatment and cDNA was amplified by PCR
 reaction. The PCR products were digested with SfiI and
 cloned into DraIII- digested pME18S-FL3 vector. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10F, by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library."
 ORIGIN
 Query Match 32.0%; Score 722; DB 4; Length 722;
 Best Local Similarity 100.0%; Pred. No. 1.2e-144;
 Matches 722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 611 ACAGCTTTCTGCCATTTTCAAAATGACCATCAGCAGATCGCGTATCAGATCTGCCAGT 670
 Db 1 ACAGCTTTCTGCCATTTTCAAAATGACCATCAGCAGATCGCGTATCAGATCTGCCAGT 60
 QY 671 CAATAAAATTTTATCATCATATAATAATTAACCCATACAGATCTGAAGCCTGAAATATTTT 730
 Db 61 CAATAAAATTTTATCATCATATAATAATTAACCCATACAGATCTGAAGCCTGAAATATTTT 120
 QY 731 TGTTTGTGAAGTCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 790
 Db 121 TGTGTGAAGTCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 QY 791 CACTGAAAAACACAGATATCAAGTTGTTGTTTGGAGTGCACCGTATGATGATGATGATGATGATGAT 850
 Db 181 CACTGAAAAACACAGATATCAAGTTGTTGTTTGGAGTGCACCGTATGATGATGATGATGATGATGAT 240

```

QY 851 ATCAGAGTACTTTGGTGTCTTACCCGGCACTACAGAGCTCCGAGGTCAATTTTGGCTTTAG 910
Db 241 ATCAGAGTACTTTGGTGTCTTACCCGGCACTACAGAGCTCCGAGGTCAATTTTGGCTTTAG 300
QY 911 GTTGTCTCAGCTTGTGATGTTTGGAGCATAGTTGATCTTATTCGAATATTAACCTTG 970
Db 301 GTTGTCTCAGCTTGTGATGTTTGGAGCATAGTTGATCTTATTCGAATATTAACCTTG 360
QY 971 GTTTCACAGTCTTTTCAGACTCATGATAGTAAAGAGCACTGGCAATGATGGAACGAATAT 1030
Db 361 GTTTCACAGTCTTTTCAGACTCATGATAGTAAAGAGCACTGGCAATGATGGAACGAATAT 420
QY 1031 TAGGACCATACCAACAACATGATTCAGAAACAAAGAAACGAAAGTATTTTCAACATA 1090
Db 421 TAGGACCATACCAACAACATGATTCAGAAACAAAGAAACGAAAGTATTTTCAACATA 480
QY 1091 ACCAGCTAGATTGGGATGAACACAGTCTCTGCTGGTAGATATGTTAGGAGACGCTGCAAC 1150
Db 481 ACCAGCTAGATTGGGATGAACACAGTCTCTGCTGGTAGATATGTTAGGAGACGCTGCAAC 540
QY 1151 CGTTGAAGGAATTTATGCTTTGTCTCATGATGAAGAACATGAGAAACTGTTTGACCTGGTTC 1210
Db 541 CGTTGAAGGAATTTATGCTTTGTCTCATGATGAAGAACATGAGAAACTGTTTGACCTGGTTC 600
QY 1211 GAAGAATGTTAGATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCATTTGACGATC 1270
Db 601 GAAGAATGTTAGATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCATTTGACGATC 660
QY 1271 CTTTCTTTGACTTATTAAGAAATGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 1330
Db 661 CTTTCTTTGACTTATTAAGAAATGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 720
QY 1331 TC 1332
Db 721 TC 722

RESULT 14
LOCUS CF110608
DEFINITION Shultzomica03859 Rat lung airway and parenchyma cDNA libraries
Rattus norvegicus cDNA clone Contig3421 5', mRNA sequence.
ACCESSION CF110608.1 GI:33166640
VERSION EST.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
REFERENCE 1. (bases 1 to 966)
AUTHORS Shultz, M.A., Zhang, L., Gu, Y.-Z., Baker, G.L., Fannuchi, M.V.,
Padua, A.M., Gurske, W.A., Morin, D., Penn, S.G., Jovanovich, S.B.,
Plopper, C.G. and Buckpitt, A.R.
TITLE Gene expression analysis in response to lung toxicants: I.
JOURNAL Sequencing and microarray development
COMMENT Am. J. Respir. Cell Mol. Biol. 30 (3), 296-310 (2004)
Contact: Shultz MA
Dept. of Molecular Biosciences, School of Veterinary Medicine
University of California, Davis
1311 Haring Hall, One Shields Avenue, Davis, CA 95616, USA
Tel: 530 752 0793
Fax: 530 752 0798
Email: ma@shultz.ucdavis.edu
Average Phred score is 20 or better. All poor quality data (Phred <
20) and vector/linker sequence has been removed.
High quality sequence stop: 966.
Location/Qualifiers
1. 966
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"

```

```

/db_xref="taxon:10116"
/clone="Contig3421"
/sex="male"
/tissue_type="airway or parenchyma"
/dev_stage="adult"
/clone_lib="Rat lung airway and parenchyma cDNA libraries"
/notice="Organ: lung; Vector: pGEM-lizf(-); Site: Eco RI;
Site 2: Not I; mRNA was isolated from microdissected rat
lung airways and parenchyma tissues."

```

ORIGIN

```

Query Match 32.0%; Score 720.8; DB 7; Length 966;
Best Local Similarity 86.0%; Pred. No. 2.2e-144;
Matches 822; Conservative 0; Mismatches 132; Indels 2; Gaps 2;
QY 678 TTTTTCATCATTAATAAATTAACCCCATACAGATCTGAAGCTCTGAATAATTTTCTTTGT 737
Db 1 TTTTTCATCATTAATAAATTAACCCCATACAGATCTGAAGCTCTGAATAATTTTCTTTGT 60
QY 738 GAAGTCTGACTATGTAGTCAAAATATAATTTTAAATGAACGTCATGATGAACGACATGAA 797
Db 61 AAAGTCTGACTATGTAGTCAAAATATAATTTTAAATGAACGTCATGATGAACGACATGAA 120
QY 798 AAACACAGATATCAAGTGTGTGACTTTGGAGTCGACAGTCATGATGAACATCAACAG 857
Db 121 AAACACAGATATCAAGTGTGTGACTTTGGAGTCGACAGTCATGATGAACATCAACAG 180
QY 858 TACTTTGTGTCTACCCGGCACTACAGAGCTCCGAGGTCAATTTTGGCTTTTGGTGTGTC 917
Db 181 TACTTTGTGTCTACCCGGCACTACAGAGCTCCGAGGTCAATTTTGGCTTTTGGTGTGTC 240
QY 918 TCAGCTTTGTGATGTTTGGAGCATAGTGTGCAATTTTATGAATATTAACCTTTGGTTCAC 977
Db 241 TCAGCTTTGTGATGTTTGGAGCATAGTGTGCAATTTTATGAATATTAACCTTTGGTTCAC 300
QY 978 AGTCTTTGAGTCTCATGATAGTAAAGAGCACCCTGGCAATGATGAACGAATATTAGGACC 1037
Db 301 AGTCTTTGAGTCTCATGATAGTAAAGAGCACCCTGGCAATGATGAACGAATATTAGGACC 360
QY 1038 CATACCAACAACATGATTTCAAGAAACAAAGAAACGCAAGTATTTTCAACCAATCAACGCT 1097
Db 361 CATCCAGCACATGATTCACAAAGACAAAGAAACGCAAGTATTTTCAACCAATCAACGCT 420
QY 1098 AGATTGGAGTGAACACAGTCTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGTTGAA 1157
Db 421 AGATTGGAGTGAACACAGTCTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGTTGAA 480
QY 1158 GGAATTTATGCTTTGTCATGATGAAGAACATGAGAAACTGTTTACCTGGTTCGAGAAAT 1217
Db 481 GGAATTTATGCTTTGTCATGATGAAGAACATGAGAAACTGTTTACCTGGTTCGAGAAAT 540
QY 1218 GTTAGAATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCAATTTGACGATCCTTTCTT 1277
Db 541 GTTAGAATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCAATTTGACGATCCTTTCTT 600
QY 1278 TGACTTTATTAAGAAAGAAATGAATGGAAATCAGTGGTCTTCTATATATATCTCTAGAA 1337
Db 601 TGACTTTATTAAGAAAGAAATGAATGGAAATCAGTGGTCTTCTATATATATCTCTAGAA 659
QY 1338 GAGATTAC- TTAAGACTGTGTGCTGCTCACTCACTCACTCACTCACTCACTCACTCACT 1396
Db 660 GCAATTACTTTTAAGACTGTGTGCTGCTCACTCACTCACTCACTCACTCACTCACTCACT 719
QY 1397 TTAATTTGTACAGTTAAGTGTAAATATGTTATGTTTGTATCAATAGCATATTAACCTTG 1456
Db 720 TTAATTTGTACAGTTAAGTGTAAATATGTTATGTTTGTATCAATAGCATATTAACCTTG 779
QY 1457 TTAAGCAAGTATGCTCTTGTATTAATGCAATAGAAATAATTAATTAATTTTCTTTTGA 1516
Db 780 TTAAGTATGTTGTTGATAATGAATGAATGCAATGATGATTAATTAATTTTCTTTGAT 839
QY 1517 ATTAACCAATTTTAAATACCTTTGAAATATCTTTTGTGTCAGTGAATTAATGATGATGATC 1576

```

Db 840 ATTGCCATTTTAAAGGCTTTACAGTCCCTTTACAGTCCCTTTGTGTCCAGTGTAAACGCTGATTGGTC 899

QY 1577 TTGCCTTTTGTACATGGAGTCACTCTGAGTGTATTTTTCAGTAAAGGAAA 1632

Db 900 TCATCTTTTGTACATTAAGGTTAACTTTGAAGTGGTTTTTTTTTCCCTCAAGTAAA 955

RESULT 15

AA631990/c

LOCUS

DEFINITION

np7405.s1 NCI CGAP Br2 Homo sapiens cDNA clone IMAGE:1132016 3' similar to SW:KCLK_MOUSE P22518 PROTEIN KINASE CLK ; , mRNA sequence.

ACCESSION

AA631990

VERSION

AA631990.1 GI:2555404

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 727)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: sgapb-remail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 2117 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 475.

FEATURES

Location/Qualifiers

1..727

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:1132016"

/sex="female, pooled"

/tissue_type="breast"

/lab_host="DH10B"

/clone_lib="NCI_CGAP_Br2"

/note="Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pooled bulk breast tumor tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T3 vector. This library is the normalized version of NCI CGAP Br1.1. Library was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 31.1%; Score 700.8; DB 1; Length 727;

Best Local Similarity 99.5%; Pred. No. 4.2e-140;

Matches 724; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 917 CTCAGCCTTTGTGATGTTGGACATAGGTTGCATTTATTGAATATTACCTTGGTTCA 976

Db 726 CTCAGCCTTTGTGATGTTGGACATAGGTTGCATTTATTGAATATTACCTTGGTTCA 668

QY 977 CAGTCTTTTCAGACTCATGATAGTAAGAGCACCTGGCAATGATGGAACGAATATTAGGAC 1036

Db 667 CAGTC-TTCAGACTCATGATAGTAAGAGCACCTGGCAATGATGGAACGAATATTAGGAC 609

QY 1037 CCATACCACACACATGATTTCAGAAAACAGAAAACGCAAGTATTTTCACCATACCCAGC 1096

Db 608 CCATACCACACACATGATTTCAGAAAACGAAACGCAAGTATTTTCACCATACCCAGC 549

QY 1097 TAGATTGGGATGAACACACAGTTCTCTGCTAGATATGTTAGGAGACGCTGCACAAACCGTTGA 1156

Db 548 TAGATTGGGATGAACACACAGTTCTCTGCTAGATATGTTAGGAGACGCTGCACAAACCGTTGA 489

QY 1157 AGGAATTTATGCTTTTGTTCATGATGAAGAACATGAGAAACCTGTTTGAACCTGGTTTGAAGAA 1216

Db 488 AGGAATTTATGCTTTTGTTCATGATGAAGAACATGAGAAACCTGTTTGAACCTGGTTTGAAGAA 429

QY 1217 TGTTAGATATGATCCAACTCAAGATTTACCTTGGATGAGCAATTCGACGATCCTTTCT 1276

Db 428 TGTTAGATATGATCCAACTCAAGATTTACCTTGGATGAGCAATTCGACGATCCTTTCT 369

QY 1277 TTGACTTTATTAAGAAAGAAATGGAATGGAATCAGTGTCTTACTATATATCTTCTAGA 1336

Db 368 TTGACTTTATTAAGAAAGAAATGGAATGGAATCAGTGTCTTACTATATATCTTCTAGA 309

QY 1337 AGAGATTACTTTAAGACTGTGTCACTCAACTAAACATTTCTAATATTTTGTAAACATTAAA 1396

Db 308 AGAGATTACTTTAAGACTGTGTCACTCAACTAAACATTTCTAATATTTTGTAAACATTAAA 249

QY 1397 TTATTTGTACAGTTAAAGTGTAAATATGTTATGTTTGTATCAATAGCATATTAACCTTG 1456

Db 248 TTATTTGTACAGTTAAAGTGTAAATATGTTATGTTTGTATCAATAGCATATTAACCTTG 189

QY 1457 TTAAGCAAGTATGCTTGTGATAATGCAATAGCAAAATTAATAATTTCTTTTGA 1516

Db 188 TTAAGCAAGTATGCTTGTGATAATGCAATAGCAAAATTAATAATTTCTTTTGA 129

QY 1517 ATTACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATTAATGTGATTGATC 1576

Db 128 ATTACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATTAATGTGATTGATC 69

QY 1577 TTGCCTTTTGTACATGGAGTCACTCTGAGTGTGATTTTTCAGTAAAGGAAATCTT 1636

Db 68 TTGCCTTTTGTACATGGAGTCACTCTGAGTGTGATTTTTCAGTAAAGGAAATCTT 9

QY 1637 GACTACTT 1644

Db 8 GACTACTT 1

Search completed: March 16, 2005, 18:40:36

Job time : 7302.99 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2005, 07:43:58 ; Search time 376.816 Seconds
(without alignments)
9796.410 Million cell updates/sec

Title: US-10-801-671-1_COPY_72_2327
Perfect score: 2256
Sequence: 1 gacactcaccagtcattt.....aaacttttgacttcagaaa 2256

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818130359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUTS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2256	100.0	2354	3	US-09-810-671-1
2	2256	100.0	2354	4	US-10-109-854-1
3	2256	100.0	2354	4	US-10-339-656-1
4	2046.6	90.7	2446	2	US-09-016-000-9
5	1178.2	52.2	1456	4	US-09-023-685-699
6	1097.4	48.6	1549	4	US-09-903-999-26
7	1097	48.6	21234	3	US-09-810-671-3
8	1097	48.6	21234	4	US-10-109-854-3
9	1097	48.6	21234	4	US-10-339-656-3
10	640.6	28.4	2254	4	US-09-919-039-238
11	481.2	21.3	1538	4	US-09-905-999-22
12	454.8	20.2	1787	4	US-09-905-999-24
13	432.4	19.2	1762	4	US-09-016-434-1439
14	430.8	19.1	1763	4	US-09-949-016-2648
15	430.8	19.1	1763	4	US-09-949-016-2649
16	430.8	19.1	1788	4	US-09-976-594-313
17	263	11.7	263	4	US-09-016-434-310
18	203.4	9.0	492	4	US-09-621-976-3125
19	185.4	8.2	479	4	US-09-621-976-3124
20	184.2	8.2	913	4	US-09-016-434-712
21	123.8	5.5	378	1	US-08-700-575-2
22	95.2	4.2	2061	2	US-08-835-170-1
23	95.2	4.2	2061	3	US-09-359-257-1
24	95.2	4.2	2061	3	US-09-371-674-1
25	95.2	4.2	2327	2	US-08-835-170-3
26	95.2	4.2	2327	3	US-09-359-257-3
27	95.2	4.2	2327	3	US-09-371-674-3

28	89	3.9	621	4	US-09-248-796A-4385	Sequence 4385, Ap
29	84.8	3.8	308	4	US-09-621-976-9631	Sequence 9631, Ap
30	81.4	3.6	260	4	US-09-313-294A-2299	Sequence 2299, Ap
31	75.6	3.4	3723	4	US-09-949-016-3590	Sequence 3590, Ap
32	75.6	3.4	71251	4	US-09-949-016-15332	Sequence 15332, A
33	68.8	3.0	282	4	US-09-313-294A-1168	Sequence 1168, Ap
34	67.2	3.0	2424	4	US-09-614-221A-518	Sequence 518, App
35	66.4	2.9	466	4	US-09-513-999C-11373	Sequence 11373, A
36	66	2.9	362	4	US-09-513-999C-14437	Sequence 14437, A
37	63.6	2.8	187169	4	US-09-949-016-12776	Sequence 12776, A
38	63.6	2.8	191569	4	US-09-949-016-15940	Sequence 15940, A
39	61.8	2.7	25882	4	US-09-949-016-14390	Sequence 14390, A
40	61.8	2.7	25882	4	US-09-949-016-14391	Sequence 14391, A
41	60.6	2.7	1141	4	US-09-806-708B-22	Sequence 22, Appl
42	60.6	2.7	59519	4	US-09-949-016-123504	Sequence 123504, A
43	60.6	2.7	119153	4	US-09-949-016-12378	Sequence 12378, A
44	60.4	2.7	19124	2	US-08-487-826B-13	Sequence 13, Appl
45	60.4	2.7	251672	4	US-09-949-016-17296	Sequence 17296, A

ALIGNMENTS

RESULT 1
US-09-810-671-1
; Sequence 1, Application US/09810671
; Patent No. 6455291
; GENERAL INFORMATION:
; APPLICANT: YAN Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758
; CURRENT APPLICATION NUMBER: US/09/810,671
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2354
; TYPE: DNA
; ORGANISM: Human
US-09-810-671-1

Query Match	100.0%	Score 2256;	DB 3;	Length 2354;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 2256;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GACACTATCCAGTCATTTTAAAGCAAGGTCCTTGAATGACGAGATTATCGGAC	60	
DB	72	GACACTATCCAGTCATTTTAAAGCAAGGTCCTTGAATGACGAGATTATCGGAC	131	
QY	61	CGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTAT	120	
DB	132	CGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTAT	191	
QY	121	CACAGAGACATTTGAAAGCGGGTATCGAATCCACTGCAGTAAATCTTCAGTCGCGACGAG	180	
DB	192	CACAGAGACATTTGAAAGCGGGTATCGAATCCACTGCAGTAAATCTTCAGTCGCGACGAG	251	
QY	181	AGAGCAGTCCTTAAAGGAAGCGCAATAGACACTGTTCAAGTCATCAGTCAGTCGGAAG	240	
DB	252	AGAGCAGTCCTTAAAGGAAGCGCAATAGACACTGTTCAAGTCATCAGTCAGTCGGAAG	311	
QY	241	AGCCACCGAAGAAAGATCCAGGATATAGAGGATGATAGGAGGGTCACCTGATCTGT	300	
DB	312	AGCCACCGAAGAAAGATCCAGGATATAGAGGATGATAGGAGGGTCACCTGATCTGT	371	
QY	301	CAAAAGTGAGACGTTCTTAAAGAGCAAGATATCAAAATCGTGGACACTTTTGGGTGAAGAGCC	360	
DB	372	CAAAAGTGAGACGTTCTTAAAGAGCAAGATATCAAAATCGTGGACACTTTTGGGTGAAGAGCC	431	
QY	361	TTTCGCAAGTTGTAGAGTCATTGATCATCGCATGGAATGCGATGTCATGTAGCAGTGA	420	

D _b	432	TTTTGGCAAAAGTTGTAGAGTGCATTGATCATCTGGCAGTCGCGATGCCATGCGNTGTACGACTGA	A91
Q _y	421	ATCCGTAAAAAATGTAGCCGGTTACCGTGGAAGCAGCTCGTTTCAGAAAATCCAAGATTTTAGAG	480
D _b	492	ATCGTAAAAAATGTAGCCGGTTACCGTGGAAGCAGCTCGTTTCAGAAAATCCAAGATTTTAGAG	551
Q _y	481	CACTTTAATAATAGTACTGATGCCAATPAGTGTCTTCCGATGTGTCAGATGCTTAGAATCGTTT	540
D _b	552	CACTTAAATAGTACTGATGCCAATPAGTGTCTTCCGATGTGTCAGATGCTTAGAATCGTTT	611
Q _y	541	GATCATCATGGTCATGTTTGTGATGTGCTTTGAACTACTGCGGACTTACTGACCTTACGATTC	600
D _b	612	GATCATCATGGTCATGTTTGTGATGTGCTTTGAACTACTGCGGACTTACTGACCTTACGATTC	671
Q _y	601	ATTTAAAGAAAACAGCTTTCTGCCATTTCAAATTTGACCACATCAGGCAGATGGCGTATCAG	660
D _b	672	ATTTAAAGAAAACAGCTTTCTGCCATTTCAAATTTGACCACATCAGGCAGATGGCGTATCAG	731
Q _y	661	ATCTGCCAGTCAATPAAATTTTTHACATCATAAATAAATTAACCCATACAGATCTGAAGCCT	720
D _b	732	ATCTGCCAGTCAATPAAATTTTTHACATCATAAATAAATTAACCCATACAGATCTGAAGCCT	791
Q _y	721	GAATAATTTTGTGTGGAAGTCTGACTATGTAGTCAAAATATAATTTCTAAAAATGAACGT	780
D _b	792	GAATAATTTTGTGTGGAAGTCTGACTATGTAGTCAAAATATAATTTCTAAAAATGAACGT	851
Q _y	781	GATGAAACGACACTGAAAAAACACACAGATATCAADAGTTCTGACCTTTGGAAGTGCACAACGTAT	840
D _b	852	GATGAAACGACACTGAAAAAACACACAGATATCAADAGTTCTGACCTTTGGAAGTGCACAACGTAT	911
Q _y	841	GATGATGAACATCACAGTACTTTGGTGTCTACCCGGCACTACAGAGCTCCCAGGTCATT	900
D _b	912	GATGATGAACATCACAGTACTTTGGTGTCTACCCGGCACTACAGAGCTCCCAGGTCATT	971
Q _y	901	TTGGCTTTAGTGTGGTCTCAGCCTTGATGTGTTTGGAGCATAGGTTGCATTCTTATGAA	960
D _b	972	TTGGCTTTAGTGTGGTCTCAGCCTTGATGTGTTTGGAGCATAGGTTGCATTCTTATGAA	1031
Q _y	961	TATTACCTTTGGTTCACAGTCTTTTTCAGACTCATGATAGTAAAGAGCACCTGGCAATGATG	1020
D _b	1032	TATTACCTTTGGTTCACAGTCTTTTTCAGACTCATGATAGTAAAGAGCACCTGGCAATGATG	1091
Q _y	1021	GAAACGAATATTAGGACCCATACCAACAACATGATTCAGAAAAACAGAAAAACGCAAGTAT	1080
D _b	1092	GAAACGAATATTAGGACCCATACCAACAACAATGATTCAGAAAAACAGAAAAACGCAAGTAT	1151
Q _y	1081	TTTCACCATAAACAGCTAGATTGGAGTGAACACAGTTTCTGCTGGTAGATATGTTAGGAGA	1140
D _b	1152	TTTCACCATAAACAGCTAGATTGGAGTGAACACAGTTTCTGCTGGTAGATATGTTAGGAGA	1211
Q _y	1141	CGCTGCAAAACCGTTTGAAGGAATTTATGCTTTTGTTCATGATGAAGAAACATGAGAAACCTGTTT	1200
D _b	1212	CGCTGCAAAACCGTTTGAAGGAATTTATGCTTTTGTTCATGATGAAGAAACATGAGAAACCTGTTT	1271
Q _y	1201	GACCTGGTTCGAAGAAATGTTAGAAATATGATCCAACTCAAAGAAATTCCTTTGGATGAAGCA	1260
D _b	1272	GACCTGGTTCGAAGAAATGTTAGAAATATGATCCAACTCAAAGAAATTCCTTTGGATGAAGCA	1331
Q _y	1261	TTGCGAGCATCCTTTCTTTTGACTTATTTAAAAAGAAATGAANTGGGAATCAGTGGTCTTTAC	1320
D _b	1332	TTGCGAGCATCCTTTCTTTTGACTTATTTAAAAAGAAATGAANTGGGAATCAGTGGTCTTTAC	1391
Q _y	1321	TATATACCTCTCTAGAGAGATTAATTAAGACTGTTCAGTCAACTAAAACATTTCTAATAT	1380
D _b	1392	TATATACCTCTCTAGAGAGATTAATTAAGACTGTTCAGTCAACTAAAACATTTCTAATAT	1451
Q _y	1381	TTTTTGTAAACATTAATTTTGTACAGTTTAAAGTGTAAATATTGTATGTTTGTATCA	1440
D _b	1452	TTTTTGTAAACATTAATTTTGTACAGTTTAAAGTGTAAATATTGTATGTTTGTATCA	1511
Q _y	1441	TAGCATAAATTAACCTTTTGAAGCAAGATPAGTCTTTGATAATGCATTAGAAAAATTTAAATTT	1500
D _b	1512	TAGCATAAATTAACCTTTTGAAGCAAGATPAGTCTTTGATAATGCATTAGAAAAATTTAAATTT	1571

Qy	1501	AA	TTTTTTCTTTTGGAAATATACGATTTTAAAATACCTTTGAAATATCCTTTGTGTCCAGTG	1566
Db	1572	AA	TTTTTCTTTTGGAAATACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCCAGTG	1631
Qy	1561	AT	AAATGTGATTCGATCTTGCCCTTTGTGACATGAGAGTCACTCTGAAGTGAATTTTTTTTGG	1620
Db	1632	AT	AAATGTGATTCGATCTTGCCCTTTGTGACATGAGAGTCACTCTGAAGTGAATTTTTTTTGG	1691
Qy	1621	AG	TAAAGGAAATCTTGACTACTTTATATCTTTAAAGGAATATCTTTATATACTTTCAAA	1680
Db	1692	AG	TAAAGGAAATCTTGACTACTTTATATCTTTAAAGGAATATCTTTATATACTTTCAAA	1751
Qy	1681	TT	TAGAACTTAACTTTTAAAGTTTTTCTTCGTGTAATGTGTGAAACGGGTGATTTATTTAA	1740
Db	1752	TT	TAGAACTTAACTTTTAAAGTTTTTCTTCGTGTAATGTGTGAAACGGGTGATTTATTTAA	1811
Qy	1741	CT	CTAGATAGCAGGTACTAGAACCAAACTCAGAAAAATGTTTACTTGTTAGAAATCTTAT	1800
Db	1812	CT	CTAGATAGCAGGTACTAGAACCAAACTCAGAAAAATGTTTACTTGTTAGAAATCTTAT	1871
Qy	1801	TAA	TTTTTAAAGTGTGTATCTTTTTCATTTGGGTGATGTCAGGGTGATAACCAAGACATTC	1860
Db	1872	TAA	TTTTTAAAGTGTGTATCTTTTTCATTTGGGTGATGTCAGGGTGATAACCAAGACATTC	1931
Qy	1861	AT	GAAAGGCAATGCAGTTTGTCCATTTGTGACAGTTTGTTTTAAATAAAACCAATACACACT	1920
Db	1932	AT	GAAAGGCAATGCAGTTTGTCCATTTGTGACAGTTTGTTTTAAATAAAACCAATACACACT	1991
Qy	1921	TT	ATTTAAAGATTAATACTTAATCTGGAAGTCAGCTTGGAAAAATGGACATTTCCAACTATG	1980
Db	1992	TT	ATTTAAAGATTAATACTTAATCTGGAAGTCAGCTTGGAAAAATGGACATTTCCAACTATG	2051
Qy	1981	TT	TGGTCAGTCACAGATATAAAATAGAAATCTTGATGAGAGGTTTCAGTTTTTAAATACC	2040
Db	2052	TT	TGGTCAGTCACAGATATAAAATAGAAATCTTGATGAGAGGTTTCAGTTTTTAAATACC	2111
Qy	2041	AA	GTCTTTAGAGTCTTAAACATTTGGCCAGCACTCTGTTTATCAAATGACATAAATACGTAA	2100
Db	2112	AA	GTCTTTAGAGTCTTAAACATTTGGCCAGCACTCTGTTTATCAAATGACATAAATACGTAA	2171
Qy	2101	AC	CTATAAGAAATTAAGTTTATTAATTTAGGCAATTTATGCTGTGATAATTTCTTTACGGGAG	2160
Db	2172	AC	CTATAAGAAATTAAGTTTATTAATTTAGGCAATTTATGCTGTGATAATTTCTTTACGGGAG	2231
Qy	2161	AA	AGAGATTTGATTTGAAAAACGATTTTGGGAAGAAAGTGCTGCTGAAAATTTCCAGAAATTT	2220
Db	2232	AA	AGAGATTTGATTTGAAAAACGATTTTGGGAAGAAAGTGCTGCTGAAAATTTCCAGAAATTT	2291
Qy	2221	AA	TGATTTGGTTACATAAACTTTTGTGACTTTCAGAAA	2256
Db	2292	AA	TGATTTGGTTACATAAACTTTTGTGACTTTCAGAAA	2327

```

RESULT 2
US-10-109-854-1
; Sequence 1, Application US/10109854
; Patent No. 6630337
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758DIV
; CURRENT APPLICATION NUMBER: US/10/109,854
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 05/810,671
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1

```

; LENGTH: 2354		Query Match		100.0%; Score 2256; DB 4; Length 2354;	
; TYPE: DNA		; ORGANISM: Homo sapien		Best Local Similarity 100.0%; Pred. No. 0;	
US-10-109-854-1		Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	GACACTCATCCAGTCATTATTTAGAGCAAGGTCCTTTGAATGACGAGGATTTATCGGAC	60		
DB	72	GACACTCATCCAGTCATTATTTAGAGCAAGGTCCTTTGAATGACGAGGATTTATCGGAC	131		
QY	61	CGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTCCCTAGACATTAT	120		
DB	132	CGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTCCCTAGACATTAT	191		
QY	121	CACAGAGACNTTGAAGCGGGTATCGAATCCACTCCAGTCAATCTTCAGTCGGCAGG	180		
DB	192	CACAGAGACNTTGAAGCGGGTATCGAATCCACTCCAGTCAATCTTCAGTCGGCAGG	251		
QY	181	AGAAGCAGTCCCTAAAGGAAGCGCAATAGACACTGTTCAAGTCATCAGTCAGTTCGAAG	240		
DB	252	AGAAGCAGTCCCTAAAGGAAGCGCAATAGACACTGTTCAAGTCATCAGTCAGTTCGAAG	311		
QY	241	AGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATAGGAGGGTCACCTGATCTGT	300		
DB	312	AGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATAGGAGGGTCACCTGATCTGT	371		
QY	301	CAAAGTGAGAGCGTTCTAAGAGCAAGATGATGAATCGTGGACACTTTGGGTGAAGAGCC	360		
DB	372	CAAAGTGAGAGCGTTCTAAGAGCAAGATGATGAATCGTGGACACTTTGGGTGAAGAGCC	431		
QY	361	TTTGGCAAGTTGTAGAGTGCAATCATCATGCAATGGATGGCATGCAATGAGCAGTGA	420		
DB	432	TTTGGCAAGTTGTAGAGTGCAATCATCATGCAATGGATGGCATGCAATGAGCAGTGA	491		
QY	421	ATCGTAAAGAAATGTAGGCGGTTACCGTGAGCAGCTCGTTTCAGAAATCCAAATATTAG	480		
DB	492	ATCGTAAAGAAATGTAGGCGGTTACCGTGAGCAGCTCGTTTCAGAAATCCAAATATTAG	551		
QY	481	CACTTAAATAGTACATCCCAATAGTGTCTCCGATGTCAGATGCCATGAGTATGGTTT	540		
DB	552	CACTTAAATAGTACATCCCAATAGTGTCTCCGATGTCAGATGCCATGAGTATGGTTT	611		
QY	541	GATCATATGTCATGTTGTATGTTGTTGAATCTGGAATCTGGAATCTGGAATCTGGA	600		
DB	612	GATCATATGTCATGTTGTATGTTGTTGAATCTGGAATCTGGAATCTGGAATCTGGA	671		
QY	601	ATTAAGAAACACAGCTTTCTGCCATTTCAAATTTGACCAATCCAGCAGATGCGGTATCAG	660		
DB	672	ATTAAGAAACACAGCTTTCTGCCATTTCAAATTTGACCAATCCAGCAGATGCGGTATCAG	731		
QY	661	ATCTGCCAGTCAATAAATTTTACATCAATAAATTTAAACCAATACAGATCTGAAGCTT	720		
DB	732	ATCTGCCAGTCAATAAATTTTACATCAATAAATTTAAACCAATACAGATCTGAAGCTT	791		
QY	721	GAATAATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	780		
DB	792	GAATAATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	851		
QY	781	GATGAACGCACACTGAAACCAACAGATATCAAGTTGTTGTTGTTGTTGTTGTTGTTGTTG	840		
DB	852	GATGAACGCACACTGAAACCAACAGATATCAAGTTGTTGTTGTTGTTGTTGTTGTTGTTG	911		
QY	841	GATGATGAACATCAAGTACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	900		
DB	912	GATGATGAACATCAAGTACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	971		
QY	901	TTGGCTTTAGGTTGCTCTCAGCCTTGTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTG	960		
DB	972	TTGGCTTTAGGTTGCTCTCAGCCTTGTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTG	1031		
QY	961	TATTACCTTGGTTTTCACAGTCTTTTCAGACTCATGATAGTAAAGAGACCTGCGCAATGATG	1020		
DB	1032	TATTACCTTGGTTTTCACAGTCTTTTCAGACTCATGATAGTAAAGAGACCTGCGCAATGATG	1091		
QY	1021	GAAAGAAATATTAGGACCCATACCAACACATGATTTCAGAAAACAGAAAACGCAAGTAT	1080		
DB	1092	GAAAGAAATATTAGGACCCATACCAACACATGATTTCAGAAAACAGAAAACGCAAGTAT	1151		
QY	1081	TTTTCACCAATAACCACTAGATTGGGATGAACACAGTCTTCTGCTGGTAGATATGTTAGGAG	1140		
DB	1152	TTTTCACCAATAACCACTAGATTGGGATGAACACAGTCTTCTGCTGGTAGATATGTTAGGAG	1211		
QY	1141	CGCTGCAACCGTTGAAGAAATTTATGCTTTGTCTGATGATGAAGACATGAGAACTGTTT	1200		
DB	1212	CGCTGCAACCGTTGAAGAAATTTATGCTTTGTCTGATGATGAAGACATGAGAACTGTTT	1271		
QY	1201	GACCTGGTTCCGAAGAAATGTTAGAAATATGATCAACTCAAGAAATTTACCTTGGATCAAGCA	1260		
DB	1272	GACCTGGTTCCGAAGAAATGTTAGAAATATGATCAACTCAAGAAATTTACCTTGGATCAAGCA	1331		
QY	1261	TTGAGACATCTCTTTCTTTGACTTATTTAAAAAGAAATGAAATGGGAATCAGTGGTCTTAC	1320		
DB	1332	TTGAGACATCTCTTTCTTTGACTTATTTAAAAAGAAATGAAATGGGAATCAGTGGTCTTAC	1391		
QY	1321	TATATACCTCTAGAGAGATTTACTTAAAGCTGTGTGATCACTCAACATTTAAATAT	1380		
DB	1392	TATATACCTCTAGAGAGATTTACTTAAAGCTGTGTGATCACTCAACATTTAAATAT	1451		
QY	1381	TTTTGTAAACATTAATTTTGTGACAGTAAAGTGAATATTTGATGTTTGTGATCA	1440		
DB	1452	TTTTGTAAACATTAATTTTGTGACAGTAAAGTGAATATTTGATGTTTGTGATCA	1511		
QY	1441	TAGCATATTAACCTTGTAAAGCAAGTATGCTTCTGATTAATGATTAAGAAATTTAAAT	1500		
DB	1512	TAGCATATTAACCTTGTAAAGCAAGTATGCTTCTGATTAATGATTAAGAAATTTAAAT	1571		
QY	1501	AATTTTCTTTTGAATTAACAATTTTAAATACCTTTGAAATATCTTTGTGTCAGTG	1560		
DB	1572	AATTTTCTTTTGAATTAACAATTTTAAATACCTTTGAAATATCTTTGTGTCAGTG	1631		
QY	1561	ATAAATGATGATGATCTTGTGCTTTGTACATGAGGCTCACCTCTGAGTGATTTTGTG	1620		
DB	1632	ATAAATGATGATGATCTTGTGCTTTGTACATGAGGCTCACCTCTGAGTGATTTTGTG	1691		
QY	1621	AGTAAAGGAAATCTTGACTACTTTTATATCTTAAAGGAAATATCTTTATATACCTCAA	1680		
DB	1692	AGTAAAGGAAATCTTGACTACTTTTATATCTTAAAGGAAATATCTTTATATACCTCAA	1751		
QY	1681	TTTAGAACCTTAACTTTTAAAGTCTTCTGTAATTTGTTGAAACGGGTGATTTATTTAA	1740		
DB	1752	TTTAGAACCTTAACTTTTAAAGTCTTCTGTAATTTGTTGAAACGGGTGATTTATTTAA	1811		
QY	1741	CTCTAGATAAGCAGTACTAGAAACCAAACTCTAGAAAATTTTACTGTTAGAAATCTAT	1800		
DB	1812	CTCTAGATAAGCAGTACTAGAAACCAAACTCTAGAAAATTTTACTGTTAGAAATCTAT	1871		
QY	1801	TAAATTTTAAAGTCTTCTGTAATTTGTTGAAACGGGTGATTTATTTAA	1860		
DB	1872	TAAATTTTAAAGTCTTCTGTAATTTGTTGAAACGGGTGATTTATTTAA	1931		
QY	1861	ATGAAAGGCAATGCAAGTTGTCATTTGCAAGTGTGTTTAAATAAACAACATACACT	1920		
DB	1932	ATGAAAGGCAATGCAAGTTGTCATTTGCAAGTGTGTTTAAATAAACAACATACACT	1991		
QY	1921	TTATTTAGATTTAAATCTTAATCTGAAAGTCACTTGGAAATGCAATTTCCAAAGTATG	1980		
DB	1992	TTATTTAGATTTAAATCTTAATCTGAAAGTCACTTGGAAATGCAATTTCCAAAGTATG	2051		
QY	1981	TTTGTGAGTCAAGATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	2040		
DB	2052	TTTGTGAGTCAAGATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	2111		
QY	2041	AAGTCTTTAGGAGTCTTAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAATACGTA	2100		


```
QY 1501 AATTTTCTTTTGAATTAACATTTTAAATACCTTTGAATATCCTTTGTGTCAGTG 1560
Db 1572 AATTTTCTTTTGAATTAACATTTTAAATACCTTTGAATATCCTTTGTGTCAGTG 1631
QY 1561 ATAATGTGATGATCTTGGCTTTTGTACATGGAGTCACCTCTGAAGTATTTTTTTG 1620
Db 1632 ATAATGTGATGATCTTGGCTTTTGTACATGGAGTCACCTCTGAAGTATTTTTTTG 1691
QY 1621 AGTAAAGGAAATCTTGACTACTTTATATCTTAAAGGAATATCTTTATATACCTCAA 1680
Db 1692 AGTAAAGGAAATCTTGACTACTTTATATCTTAAAGGAATATCTTTATATACCTCAA 1751
QY 1681 TTTAGAACCTTAACCTTTAAAGTTTCTTCTGTAATTTGTGAACGGGTGATTTATTA 1740
Db 1752 TTTAGAACCTTAACCTTTAAAGTTTCTTCTGTAATTTGTGAACGGGTGATTTATTA 1811
QY 1741 CTCAGATAACGAGTACTAGAAACCAAACTCAGAAATCTTTACTGTTAGATTTCTAT 1800
Db 1812 CTCAGATAACGAGTACTAGAAACCAAACTCAGAAATCTTTACTGTTAGATTTCTAT 1871
QY 1801 TAAATTTTAAAGTGTGATTTCTTTTCTTGGGTGATGTCAGGGTGATAACAGACATTC 1860
Db 1872 TAAATTTTAAAGTGTGATTTCTTTTCTTGGGTGATGTCAGGGTGATAACAGACATTC 1931
QY 1861 ATGAAAGGCAATGCAAGTTTGTCCATTTGACAGTTTGTATTAATAAACACATACACT 1920
Db 1932 ATGAAAGGCAATGCAAGTTTGTCCATTTGACAGTTTGTATTAATAAACACATACACT 1991
QY 1921 TTATTTAGATTAATCTTAAGTGAAGTGCCTTGGAAATGAGACATTTTCCAAGTATG 1980
Db 1992 TTATTTAGATTAATCTTAAGTGAAGTGCCTTGGAAATGAGACATTTTCCAAGTATG 2051
QY 1981 TTTGTGAGTCACAGATATAAAATAGAAATCTCATGAGAGGTTTCAGTTTTTAAATACC 2040
Db 2052 TTTGTGAGTCACAGATATAAAATAGAAATCTCATGAGAGGTTTCAGTTTTTAAATACC 2111
QY 2041 AAGTCCTTAGGAGTCTTAACATTTGCCAGCATCTGTTTATCAATGACATAAATACGTA 2100
Db 2112 AAGTCCTTAGGAGTCTTAACATTTGCCAGCATCTGTTTATCAATGACATAAATACGTA 2171
QY 2101 ACCTATAGAAATTAAGTTTATTAATAGGCAATTTATGTCGTGATTAATCTTACGGGAG 2160
Db 2172 ACCTATAGAAATTAAGTTTATTAATAGGCAATTTATGTCGTGATTAATCTTACGGGAG 2231
QY 2161 AAAGAGGATTTGATTTGAAAGCAGTTTGGGAAGAGTGTGCTGAAATTTTCCAGAAATTT 2220
Db 2232 AAAGAGGATTTGATTTGAAAGCAGTTTGGGAAGAGTGTGCTGAAATTTTCCAGAAATTT 2291
QY 2221 AATTGATTTGTTACATAAACTTTTGTGACTTCAGAAA 2256
Db 2292 AATTGATTTGTTACATAAACTTTTGTGACTTCAGAAA 2327
```

RESULT 4
US-09-016-000-9
; Sequence 9, Application US/09016000
; Patent No. 5962232
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Bandman, Olga
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; TITLE OF INVENTION: PROTEIN KINASE MOLECULES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA

ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,000
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0465 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2446 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: NEUTPMT01
CLONE: 339963
US-09-016-000-9

Query Match 90.7%; Score 2046.6; DB 2; Length 2446;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 2149; Conservative 0; Mismatches 4; Indels 90; Gaps 1;

QY 14 GTCATTATTTAGAACGAAGTCTCTTGAATGACGAGATTTATCGGACCGGAGATACGTTG 73
Db 288 GTCATTATTTAGAACGAAGTCTCTTGAATGACGAGATTTATCGGACCGGAGATACGTTG 347

QY 74 ACGATACAGGAATGACTACTGTGAAGATATGTTCTTAGACATTTATCACAGACATTTG 133
Db 348 ACGAATACAGGAATGACTACTGTGAAGATATGTTCTTAGACATTTATCACAGACATTTG 407

QY 134 AAAGCGGTATCGAATCCACTGCGAGTAACTCTCAGTCCGACGAGGAGAGACGCTCCTA 193
Db 408 AAAGCGGTATCGAATCCACTGCGAGTAACTCTCAGTCCGACGAGGAGAGACGCTCCTA 467

QY 194 AAAGGAAGCGCAATAGACACTGTTTCAAGTCAATCAGTCAAGTTCGAAAGAGCCACCGAAGGA 253
Db 468 AAAGGAAGCGCAATAGACACTGTTTCAAGTCAATCAGTCAAGTTCGAAAGAGCCACCGAAGGA 515

QY 254 AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAGTGGAGAGG 313
Db 516 ----- 515

QY 314 TTCTAAGAGCAAGATATGAAATCGTGGACACTTTTGGGTGAAGGAGCCCTTTGSCAAAGTTG 373
Db 516 -----AATCCGTGGACACTTTTGGGTGAAGGAGCCCTTTGSCAAAGTTG 557

QY 374 TAGAGTGCATTTGATCATGGCATGATGGCATGTCATGTAGCAGTGAATAATCGTAAATAATG 433
Db 558 TAGAGTGCATTTGATCATGGCATGATGGCATGTCATGTAGCAGTGAATAATCGTAAATAATG 617

QY 434 TAGGCGGTTACCGTGAAGCAGCTGTTTCAAGTAAATCCAAAGTATTTAGAGCAGCTTAATAATGA 493
Db 618 TAGGCGGTTACCGTGAAGCAGCTGTTTCAAGTAAATCCAAAGTATTTAGAGCAGCTTAATAATGA 677

QY 494 CTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGTGTAGATGTTTTCATCATCATGTGTC 553
Db 678 CTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGTGTAGATGTTTTCATCATCATGTGTC 737

QY 554 ATGTTTGTATTGTTGTTTGAACCTACTGGGACTTAGTACTTACGATTTTTCATTAAAGAAACA 613

Db ATGTTGTTATGTTGTTGAACTACTGGGACTAGTACTTACGATTTCAATTAAGAAACA 797
QY GTTTCCTGCCATTTCAAAATGACCACTACAGGAGATGGCGTATCAGATCTGCCAGTCAA 673
Db GTTTCCTGCCATTTCAAAATGACCACTACAGGAGATGGCGTATCAGATCTGCCAGTCAA 857
QY TAAATTTTTTACATCATATAATAATTAACCCATACAGATCTGAAGCTCGAAATATTTTGT 733
Db TAAATTTTTTACATCATATAATAATTAACCCATACAGATCTGAAGCTCGAAATATTTTGT 917
QY TTGTGAAGTCTGACTATGATGCTCAATATAATTTCAAAATGAACGCTGATGAACACAC 793
Db TTGTGAAGTCTGACTATGATGCTCAATATAATTTCAAAATGAACGCTGATGAACACAC 977
QY TGAATAACACAGATATCAAAAGTTGTGACTTTGGAGTGGCAACGATGATGAACATC 853
Db TGAATAACACAGATATCAAAAGTTGTGACTTTGGAGTGGCAACGATGATGAACATC 1037
QY ACAGTACTTTGGTGTCTACCGGCACTACAGAGCTCCGAGGTCAATTTGGCTTTAGGTT 913
Db ACAGTACTTTGGTGTCTACCGGCACTACAGAGCTCCGAGGTCAATTTGGCTTTAGGTT 1097
QY GGTCTCAGCCTTTGATGTTTGGAGCATAGGTGTCATTTTATTGAATATTACCTTGGTT 973
Db GGTCTCAGCCTTTGATGTTTGGAGCATAGGTGTCATTTTATTGAATATTACCTTGGTT 1157
QY TCACAGTCTTTTCAGACTCATGATAGTAAGACCTGGCAATGATGAACGATATTAG 1033
Db TCACAGTCTTTTCAGACTCATGATAGTAAGACCTGGCAATGATGAACGATATTAG 1217
QY GACCCATACCAACACACATGATTTCAGAAAAACGAAAGTATTTCACCAATACC 1093
Db GACCCATACCAACACACATGATTTCAGAAAAACGAAAGTATTTCACCAATACC 1277
QY AGCTAGATTGGGATGAACACAGTCTCTGCTGGTATGATGTTAGGAGCGCTGCAACCGT 1153
Db AGCTAGATTGGGATGAACACAGTCTCTGCTGGTATGATGTTAGGAGCGCTGCAACCGT 1337
QY TGAAGGAATTTATGCTTTGTCATGATGAAGACATGAGAACTGTTGACCTGGTTCGAA 1213
Db TGAAGGAATTTATGCTTTGTCATGATGAAGACATGAGAACTGTTGACCTGGTTCGAA 1397
QY GAATGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGACGATCCT 1273
Db GAAATGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGACGATCCT 1457
QY TCTTTGACTTTATTAAGAAATGAATGGGAATCAGTGGTCTTACTATATCTCTCT 1333
Db TCTTTGACTTTATTAAGAAATGAATGGGAATCAGTGGTCTTACTATATCTCTCT 1517
QY AGAGAGATTACTTAAGACTGTGTCAGTCACTAAACATTTCTAATTTTGTAAACATT 1393
Db AGAGAGATTACTTAAGACTGTGTCAGTCACTAAACATTTCTAATTTTGTAAACATT 1577
QY AAATTAATTTGTACAGTTAAAGTAAATATGATGTTTGTATCAATAGCATATTAAC 1453
Db AAATTAATTTGTACAGTTAAAGTAAATATGATGTTTGTATCAATAGCATATTAAC 1637
QY TTGTTAAGCAAGTATGTTCTTGATATGCAATAGAAAAATTAATAATTTTCTTTTT 1513
Db TTGTTAAGCAAGTATGTTCTTGATATGCAATAGAAAAATTAATAATTTTCTTTTT 1697
QY GAAATTAACATTTTAAATACCTTTGAATATCCTTTGTGTCAGTGAATGATGATTG 1573
Db GAAATTAACATTTTAAATACCTTTGAATATCCTTTGTGTCAGTGAATGATGATTG 1757
QY ATCTTGCTTTTGTACATGGAGGTCACTCTGAAAGTATTTTTTTTGTAGTAAAGAAAT 1633
Db ATCTTGCTTTTGTACATGGAGGTCACTCTGAAAGTATTTTTTTTGTAGTAAAGAAAT 1817
QY CTGTGACTATTTATATCTTAAAGGAATATCTTTATATATCTTCAATTTAGAACTTAAC 1693

Db CTTGACTACTTTATATTTCTTAAAGGAATATTTCTTTATATATCTTCAAAATTTAGAACTTAAC 1877
QY TTTTAAAGTTTTTCTCTCTGTAATTTGTTGAACGGGTGATTTATTTAACTCTAGATAAGCA 1753
Db TTTTAAAGTTTTTCTCTCTGTAATTTGTTGAACGGGTGATTTATTTAACTCTAGATAAGCA 1937
QY GGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGAAATTTCTATTTAAATTTTAAAGT 1813
Db GGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGAAATTTCTATTTAAATTTTAAAGT 1997
QY TTGTAATCTTTTTCATTTGGGTGATGTCAGGGTGAATAACCAACATTCATGGAAGGCGATG 1873
Db TTGTAATCTTTTTCATTTGGGTGATGTCAGGGTGAATAACCAACATTCATGGAAGGCGATG 2057
QY CAGTTTGTCCATTTGTCAGAGTTTCTGTTTAAATAAAACCATACACACTTTTATTAAAGATTA 1933
Db CAGTTTGTCCATTTGTCAGAGTTTCTGTTTAAATAAAACCATACACACTTTTATTAAAGATTA 2117
QY AAATCTAACTGGAAGTCAAGTTTGGAAAAATGGAATTTTCCAAAGTATGTTTGGTGAATC 1993
Db AAATCTAACTGGAAGTCAAGTTTGGAAAAATGGAATTTTCCAAAGTATGTTTGGTGAATC 2177
QY AGATAATAAATAGAAATTTCTGATGAGAGGTTTCAGTTTTTAAACCAAGTCTTACGAG 2053
Db AGATAATAAATAGAAATTTCTGATGAGAGGTTTCAGTTTTTAAACCAAGTCTTACGAG 2237
QY TCTTAAACATTTGCCAGCATCTGTTTATCAAAATGACATAAATACCTATAAAGAAAT 2113
Db TCTTAAACATTTGCCAGCATCTGTTTATCAAAATGACATAAATACCTATAAAGAAAT 2297
QY AAGTTTATTAATAGGCAATTTATGTCGTGATTAATTTCTTACGGGAGAAAGAGGATTTGA 2173
Db AAGTTTATTAATAGGCAATTTATGTCGTGATTAATTTCTTACGGGAGAAAGAGGATTTGA 2357
QY TTGAAAGCAGTTTGGGAAGAGTCTGCTGTAATTTCCAGAAATTTCCAGAAATTTAATTTGTTGTTA 2233
Db TTGAAAGCAGTTTGGGAAGAGTCTGCTGTAATTTCCAGAAATTTCCAGAAATTTAATTTGTTGTTA 2417
QY CATAACTTTTGTGACTTTCAGAAA 2256
Db CATAACTTTTGTGACTTTCAGAAA 2440

RESULT 5

US-09-023-655-699
; Sequence 699. Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: COMPOSITION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023.655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
 NAME: Zeller, Karen J.
 REGISTRATION NUMBER: 37,071
 REFERENCE/DOCKET NUMBER: PA-0001 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (650) 855-0555
 TELEFAX: (650) 845-4166
 INFORMATION FOR SEQ ID NO: 699:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1456 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: HNT2AGT01
 CLONE: 488842
 US-09-023-655-699

Query Match 52.2%; Score 1178.2; DB 4; Length 1456;

Best Local Similarity 99.7%; Pred. No. 6.7e-299;

Matches 1191; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 932 TTGGAGCATAGGTTGCATCTCTTATTGAATATTACCTTGGTTTCACAGTCTTTCAGACTC 991
 DB 262 TGTAGCGCATAGGTTGCATCTCTTATTGAATATTACCTTGGTTTCACAGTCTTTCAGACTC 321
 QY 992 ATGATAGTAAAGAGCACCTGGCAATGATGGAACGAATATTAGGACCCATACACAAACACA 1051
 DB 322 ATGATAGTAAAGAGCACCTGGCAATGATGGAACGAATATTAGGACCCATACACAAACACA 381
 QY 1052 TGATTGAGAAACAAAGAAACGCAAGTATTTTCCACATAACCCAGCTAGATTGGGATGAAC 1111
 DB 382 TGATTGAGAAACAAAGAAACGCAAGTATTTTCCACATAACCCAGCTAGATTGGGATGAAC 441
 QY 1112 ACAGTTCTGCTGGTAGATATGTTAGGAGCGCTGCAAAACCGTTGAAGGAATTTATGCTTT 1171
 DB 442 ACAGTTCTGCTGGTAGATATGTTAGGAGCGCTGCAAAACCGTTGAAGGAATTTATGCTTT 501
 QY 1172 GTCATGATGAACATCAGAAACGTTTGACCTGGTTCGGAAGTATTTAGAAATGATC 1231
 DB 502 GTCATGATGAACATCAGAAACGTTTGACCTGGTTCGGAAGTATTTAGAAATGATC 561
 QY 1232 CAACCTCAAGAAATTAACCTTGAATGAAGCATTCGACATCCTTTCTTTGACTTATTAAAAA 1291
 DB 562 CAACCTCAAGAAATTAACCTTGAATGAAGCATTCGACATCCTTTCTTTGACTTATTAAAAA 621
 QY 1292 AGAAATGAATGGGAATCAGTGGTCTTACTATATATCTTCTAGAGAGATTTACTTAAGA 1351
 DB 622 AGAAATGAATGGGAATCAGTGGTCTTACTATATATCTTCTAGAGAGATTTACTTAAGA 681
 QY 1352 CTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTTAAATTTTGTACAGTT 1411
 DB 682 CTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATTTAAATTTTGTACAGTT 741
 QY 1412 AAGTGAATATTTGATGTTTGTATCAATAGCATAATTAACCTTGAAGCAAGTATGGT 1471
 DB 742 AAGTGAATATTTGATGTTTGTATCAATAGCATAATTAACCTTGAAGCAAGTATGGT 801
 QY 1472 CTTGATAATGCAATTAAGAAAAATTAATTTTCTTTTCTTTTGAATTA-CCATTTTAA 1530
 DB 802 CTTGATAATGCAATTAAGAAAAATTAATTTTCTTTTCTTTTGAATTAATCCATTTTAA 861
 QY 1531 ATACCTTTGAATATCTTTGTCAGTGAATTAATGATGATCTTGCCTTTGTACA 1590
 DB 862 ATACCTTTGAATATCTTTGTCAGTGAATTAATGATGATCTTGCCTTTGTACA 921
 QY 1591 TGGAGGTCACTCTCAAGTGATTTTCTTTTGTAGTAAAGAAATCTTGACTACTTTATATT 1650
 DB 922 TGGAGGTCACTCTCAAGTGATTTTCTTTTGTAGTAAAGAAATCTTGACTACTTTATATT 981
 QY 1651 CTTAAAGGAATATCTTTATATATCTTCAAAATTTAGAACTTAATTTAAAGTTTCTTC 1710

DB 982 CTTAAAGGAATATCTTTATATATATCTTCAAAATTTAGAACTTAACCTTTAAAGTTTCTTC 1041
 QY 1711 TGTAAATTTGAAACGGGTGATTTATTATTAACTCTAGATAAGCAGGTACTAGAAACCAAAA 1770
 DB 1042 TGTAAATTTGAAACGGGTGATTTATTATTAACTCTAGATAAGCAGGTACTAGAAACCAAAA 1101
 QY 1771 CTCAGAAATGTTTACTGTTAGAAATCTTATTAAATTTTAAAGTTTGTATTCTTTTTCATT 1830
 DB 1102 CTCAGAAATGTTTACTGTTAGAAATCTTATTAAATTTTAAAGTTTGTATTCTTTTTCATT 1161
 QY 1831 GGGTATGTCAGGGTGATAACACAGACATTCATGGAAGGCATGAGTTTGTCCATTGTGA 1890
 DB 1162 GGGTATGTCAGGGTGATAACACAGACATTCATGGAAGGCATGAGTTTGTCCATTGTGA 1221
 QY 1891 CAGTTTGTTTAATAAAACACATACACACTTTTATTAAAGTTTAAATCTAACTGGAAGT 1950
 DB 1222 CAGTTTGTTTAATAAAACACATACACACTTTTATTAAAGTTTAAATCTAACTGGAAGT 1281
 QY 1951 CAGTTTGGAAAAATGGACATTTCCCAAGTATGTTTGGTGAATCAGATATATAAATAGAAA 2010
 DB 1282 CAGCTTGGAAAAATGGACATTTCCCAAGTATGTTTGGTGAATCAGATATATAAATAGAAA 1341
 QY 2011 TTCTGATCAGAGGTTTTCAGTTTAAATACCAAGTCCTTAGGAGTCTTAACATTGGCCAGC 2070
 DB 1342 TTCTGATGAGAGGTTTTCAGTTTAAATACCAAGTCCTTAGGAGTCTTAACATTGGCCAGC 1401
 QY 2071 ATCTGTTTATCAATGACATAAATACGTAACCTATAAGAAATTAAGTTTATTAAAT 2125
 DB 1402 ATCTGTTTATCAATGACATAAATACGTAACCTATAAGAAATTAAGTTTATTAAAT 1456

RESULT 6

US-09-905-999-26

Sequence 26, Application US/09905999

Patent No. 6797513

GENERAL INFORMATION:

APPLICANT: ULLRICH, Axel

APPLICANT: NAVIER, Oliver

TITLE OF INVENTION: CLK PROTEIN KINASES AND RELATED PRODUCTS AND METHODS

FILE REFERENCE: 038602/0431

CURRENT APPLICATION NUMBER: US/09/905,999

PRIOR FILING DATE: 2001-07-17

PRIOR APPLICATION NUMBER: 09/127,248

PRIOR FILING DATE: 1999-07-31

PRIOR APPLICATION NUMBER: PCT/IB97/00946

PRIOR FILING DATE: 1997-06-17

PRIOR APPLICATION NUMBER: US 08/877,150

PRIOR FILING DATE: 1997-06-17

PRIOR APPLICATION NUMBER: US 60/034,286

PRIOR FILING DATE: 1996-12-19

NUMBER OF SEQ ID NOS: 26

SOFTWARE: PatentIn version 3.0

SEQ ID NO 26

LENGTH: 1549

TYPE: DNA

ORGANISM: Mus musculus

US-09-905-999-26

Query Match

48.6%; Score 1097.4; DB 4; Length 1549;

Best Local Similarity 90.6%; Pred. No. 1.1e-277;

Matches 1170; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 14 GTCATTATTTAGAACAGGTCCTTGAATGACGAGATTTATCGGACCGGAGATACGTTG 73
 DB 239 GTCATTATTTAGAACAGATGCTTGAATGAGAGATTTATCGGACCGGAGATACATTG 298
 QY 74 ACGAATACAGGAATGACTTCTGTGAAGGATATGTTCTTAGACATTATCA CAGAGACATG 133
 DB 299 ATGAATACAGAAATGACTTACTGCGAAGGATATGTTCCAAAGACATTACCATGAGACGTTG 358
 QY 134 AAAGCGGTATCGAATCCACTGCGAGTAAATCTTCACTCGGACGAGGAGAGACGTCCTA 193
 DB 359 AAAGCACTTACCGGATCCATTGCGAGTAAATCCTAGTCAGGAGCAGGAGACGCCCTA 418

QY 194 AAAGGAGCGCAATAGACACTGTTCAAGTCATCAGTCACGTTTCGAAGAGCCACCGAAGGA 253
DB 419 AGAGAAAGCGTAATAGACCCCTGTGCAAGTCATCAGTCGCATTCGAGAGCCACCGAAGGA 478
QY 254 AAAGATCCAGAGGATATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAAGTGAGAG 313
DB 479 AAAGATCCAGAGGATATAGAGGATGATGAGGAGGTCACCTGATCTGTCAAAAGTGAGAG 538
QY 314 TTCTAAGAGCAAGATATGAATCTGGACACTTTGGGTGAAGGAGCCCTTGGCAAGTTG 373
DB 539 TTCTAAGAGCAAGATATGAATCTGGACACTTTAGGTGAAGGAGCCCTTGGCAAGTTG 598
QY 374 TAGAGTGCAATGATCATCGCATGATGCGCATGTAGCAGTGAATAATCTGTAATAAATG 433
DB 599 TAGAGTGCAATGATCATCGCATGATGCGCATGTAGCAGTGAATAATCTGTAATAAATG 658
QY 434 TAGGCGGTTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAAGAGCATTAAATAGTA 493
DB 659 TAGGAGGTTACCGGAGGAGCAGCTCGTTCAGAAATCCAAAGTATTAAGAGCATTGAACAGCA 718
QY 494 CTGATCCCAATAGTGTCTTCGATGTGTCCAGATGCTAGAAATGTTTGCATCATCATGTC 553
DB 719 CTGATCCCAATAGTGTCTTCGATGTGTCCAGATGCTAGAAATGTTTGCATCATCATGTC 778
QY 554 ATGTTTGTATGTTGTTGAACTACTGGGACTTAGTACTTACGATTTCAATTAAGAAACA 613
DB 779 ATGTTTGTATGTTGTTGAGCTGTGGGACTTAGTACTTACGATTTCAATTAAGAAACA 838
QY 614 GCTTTCTGCCATTTCAAATGACCAATCAGCAGATGCGGTATCAGATCTGCCAGTCAA 673
DB 839 GTTTTCTGCCATTTCAAATGATCACATCAGCAGATGCGGTATCAGATCTGCCAGTCAA 898
QY 674 TAAATTTTACATCATATAATTAATTAACCCATACAGATCTGAAAGCTGAAATATTTTGT 733
DB 899 TAAATTTTACATCATATAATTAATTAACCCATACAGATCTGAAAGCTGAAATATTTTAT 958
QY 734 TTGTGAAGTCTGACTATGCTAGTCAAAATATAATTTCAAAATGAAACGCTGATGAACGACAC 793
DB 959 TTGTGAAGTCTGACTATGCTAGTCAAAATATAATTTCAAAATGAAACGCTGATGAACGACAC 1018
QY 794 TGAATAACACAGATATCAAAATGTTGTGACTTTGGAAGTGCACGATGATGATGAACATC 853
DB 1019 TGAATAACACAGATATCAAAATGTTGTGACTTTGGAAGTGCACGATGATGATGAACATC 1078
QY 854 ACAGTACTTTGGTGTCTACCGGCACTACAGAGTCCCGAGGTCAATTTTGGCTTTAGGT 913
DB 1079 ATAGTACTTTGGTGTCTACCGGCACTACAGAGTCCCGAGGTCAATTTTGGCTTTAGGT 1138
QY 914 GGTCTCAGCCTTTGATGTTTGGAGCATAGGTGTCATTTCTTATTGAAATATTACCTTGGTT 973
DB 1139 GGTCTCAGCCTTTGATGTTTGGAGCATAGGTGTCATTTCTTATTGAGTACTACCTTGGT 1198
QY 974 TCACAGTCTTTCAGACTCATATAGTAAAGAGCAGCTGGCAATGATGAAACGAAATATPAG 1033
DB 1199 TCACAGTCTTTCAGAGCCACGATAGTAAAGAGCAGCTGGCAATGATGAAAGGATCTTAG 1258
QY 1034 GACCCATACCAACACATGATTCAGAAACCAAGAAACGCAAGTATTTTCAACCATACC 1093
DB 1259 GACCCATACCAACATGATTCAGAAACCAAGAAACGCAAGTATTTTCAACCATACC 1318
QY 1094 AGCTAGATTGGATGAACACAGTCTCTGCTGTGTAGATATGTTAGGAGAGCGCTGCAACCGT 1153
DB 1319 AGCTAGATTGGAGCAGCATAGTTCAGCTGGAGATATGTTAGGAGAGCGCTGCAACCGT 1378
QY 1154 TGAAGGAATTTATGCTTTGTCATGATGAAGAACATGAGAAACATGTTTGCCTGTTGCA 1213
DB 1379 TAAAGGAATTTATGCTGTGTATGACGAGAGCAGTGAAGTGTGTTGACCTGTTGCA 1438
QY 1214 GAATGTTAGATATGATCCAACTCAAGAAATTTACCTGAGTGAAGCATTTGAGCATCTT 1273
DB 1439 GAATGTTGAGATATGACCCAGGAGAGGATCACCTTGGATGAAGCATTTGAGCATCTT 1498

QY 1274 TCTTTGACTTATTAAAAAGAAATGAATGG 1304
DB 1499 TCTTTGACTTATTAAAAAGAAATGAGTGG 1529
RESULT 7
US-09-810-671-3
; Sequence 3, Application US/09810671
; Patent No. 6455291
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000758
; CURRENT APPLICATION NUMBER: US/09/810,671
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 21234
; TYPE: DNA
; ORGANISM: Human
US-09-810-671-3
Query Match 48.6%; Score 1097; DB 3; Length 21234;
Best Local Similarity 100.0%; Pred. No. 5, 4e-277;
Matches 1097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1157 AGGAATTTATGCTTTGTGCATGATGAAGAACATGAGAAACTGTTTGACCTGGTTGGAAGAA 1216
DB 18138 AGGAATTTATGCTTTGTGCATGATGAAGAACATGAGAAACTGTTTGACCTGGTTGGAAGAA 18197
QY 1217 TGTTAGATATGATCCCACTCAAGAAATTAACCTTGGATGAAGCATGAGCATCCTTTCT 1276
DB 18198 TGTTAGATATGATCCCACTCAAGAAATTAACCTTGGATGAAGCATGAGCATCCTTTCT 18257
QY 1277 TTGACTTTATTAAGAAAGAAATGGAATCGAGTCTTACTATATCTTCTCTAGA 1336
DB 18258 TTGACTTTATTAAGAAAGAAATGGAATCGAGTCTTACTATATCTTCTCTAGA 18317
QY 1337 AGAGATTTCTTAAGACTGTGTCACTCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1396
DB 18318 AGAGATTTCTTAAGACTGTGTCACTCAATCAATCAATCAATCAATCAATCAATCAATCAAT 18377
QY 1397 TTATTTGTACAGTTAAGTGTAAATATTTGATGTTTGTATCAATAGCATATTAATCACTG 1456
DB 18378 TTATTTGTACAGTTAAGTGTAAATATTTGATGTTTGTATCAATAGCATATTAATCACTG 18437
QY 1457 TTAAGCAAGTATGCTCTTGATAATGCAATAGAAAAATTAATTTTCTTTTCAA 1516
DB 18438 TTAAGCAAGTATGCTCTTGATAATGCAATAGAAAAATTAATTTTCTTTTCAA 18497
QY 1517 ATTACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGAATAATGATGATGATC 1576
DB 18498 ATTACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGAATAATGATGATGATC 18557
QY 1577 TTGCTTTTGTACATGGAGTCACTCTGAAGTGAATTTTGTAGTAAAAAGGAATCTT 1636
DB 18558 TTGCTTTTGTACATGGAGTCACTCTGAAGTGAATTTTGTAGTAAAAAGGAATCTT 18617
QY 1637 GACTACTTTTATTTCTTAAAGGAATATTTCTTTATATCTTCAAAATTTAGAACTTAACTTT 1696
DB 18618 GACTACTTTTATTTCTTAAAGGAATATTTCTTTATATCTTCAAAATTTAGAACTTAACTTT 18677
QY 1697 AAAAGTTTTTCTCTGTAATTTGTGAACGGGTGATTTATTAATCTTAGATTAAGCAGT 1756
DB 18678 AAAAGTTTTTCTCTGTAATTTGTGAACGGGTGATTTATTAATCTTAGATTAAGCAGT 18737
QY 1757 ACTAGAACCAAACTCAGAAAAATGTTTACGTTGATGATTTCTATTAATTTTAACTGTTG 1816
DB 18738 ACTAGAACCAAACTCAGAAAAATGTTTACGTTGATGATTTCTATTAATTTTAACTGTTG 18797

QY 1817 TATTCCTTTTTCATCGGGTGATCTCAGGGTGATAACACAGACATTCATCGAAAGGCATGCAG 1876
Db 18798 TATTCCTTTTTCATCGGGTGATCTCAGGGTGATAACACAGACATTCATCGAAAGGCATGCAG 18857
QY 1877 TTTGTCCATTGACAGTTTGTATTAATAAACCACATACACATTTTATTTAAGATTAAAA 1936
Db 18856 TTTGTCCATTGACAGTTTGTATTAATAAACCACATACACATTTTATTTAAGATTAAAA 18917
QY 1937 TCTAACTGGAAAGTCAGCTTGGAAATGGACATTTCCAAAGTATGTTTGGTGAGTCACACAGA 1996
Db 18918 TCTAACTGGAAAGTCAGCTTGGAAATGGACATTTCCAAAGTATGTTTGGTGAGTCACACAGA 18977
QY 1997 TATAAAATAGAAATCTGATGAGAGTTTCAGTTTAAATACCAAGTCCTTACGAGTCT 2056
Db 18978 TATAAAATAGAAATCTGATGAGAGTTTCAGTTTAAATACCAAGTCCTTACGAGTCT 19037
QY 2057 TAACTTCGCCAGCATCTGTTTATCAATGACATTAATACGTAACCTTATAGAAATTAAG 2116
Db 19038 TAACTTCGCCAGCATCTGTTTATCAATGACATTAATACGTAACCTTATAGAAATTAAG 19097
QY 2117 TTTTAAATAGGCAATTTATGCTGTGATATTTCTTACGGGAGAAAGAGGATTTGATTG 2176
Db 19098 TTTTAAATAGGCAATTTATGCTGTGATATTTCTTACGGGAGAAAGAGGATTTGATTG 19157
QY 2177 GAAAGCAGTTTGGGAGAAAGTCTGCTGAAATTTCCAGAAATTTAAATGATTGTTACAT 2236
Db 19158 GAAAGCAGTTTGGGAGAAAGTCTGCTGAAATTTCCAGAAATTTAAATGATTGTTACAT 19217
QY 2237 AAACCTTTTGTACTTCAG 2253
Db 19218 AAACCTTTTGTACTTCAG 19234

RESULT 8

US-10-109-854-3
; Sequence 3, Application US/10109854
; Patent No. 6630337
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758DIV
; CURRENT APPLICATION NUMBER: US/10/109,854
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 21234
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-109-854-3

Query Match 48.6%; Score 1097; DB 4; Length 21234;
Best Local Similarity 100.0%; Pred. No. 5.4e-277;
Matches 1097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1157 AGGAATTTATGCTTGTGATGATGAAGAACATGAGAAACTGTTTACCTGGTTCGAAGAA 1216
Db 18138 AGGAATTTATGCTTGTGATGATGAAGAACATGAGAAACTGTTTACCTGGTTCGAAGAA 18197
QY 1217 TGTTAGATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCAATTCAGCATCCTTTCT 1276
Db 18198 TGTTAGATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCAATTCAGCATCCTTTCT 18257
QY 1277 TTGACTTTATTAAGAAAGAAATGGAATGGAATCACTGCTTACTATATATCTTCTTCTAGA 1336
Db 18258 TTGACTTTATTAAGAAAGAAATGGAATGGAATCACTGCTTACTATATATCTTCTTCTAGA 18317

RESULT 9

US-10-339-656-3
; Sequence 3, Application US/10339656
; Patent No. 6733978
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

QY 1337 AGAGATTACTTAAGACTGTGTCAAGTCAACTAAACATTTCTAAATTTTGTAAACATTAAA 1396
Db 18318 AGAGATTACTTAAGACTGTGTCAAGTCAACTAAACATTTCTAAATTTTGTAAACATTAAA 18377
QY 1397 TTTTGTGACAGTTTAAAGTGAATAATTTATGATGTTTGTATCAATAGCATTAATTAACCTG 1456
Db 18378 TTTTGTGACAGTTTAAAGTGAATAATTTATGATGTTTGTATCAATAGCATTAATTAACCTG 18437
QY 1457 TTAAGCAAGTATGGTCTTGATAATGCAATAGAAAAATTTAAAAATTTCTTTTCTTTTGA 1516
Db 18438 TTAAGCAAGTATGGTCTTGATAATGCAATAGAAAAATTTAAAAATTTCTTTTCTTTTGA 18497
QY 1517 ATTACCAATTTTAAATACCTTTTGAATATCTTTGTGTCCAGTGATAAATGATGATGATC 1576
Db 18498 ATTACCAATTTTAAATACCTTTTGAATATCTTTGTGTCCAGTGATAAATGATGATGATC 18557
QY 1577 TTGCTTTTGTACATGGAGGTCACTCTGGAAGTGAATTTTGTGATTAAGGAATCTT 1636
Db 18558 TTGCTTTTGTACATGGAGGTCACTCTGGAAGTGAATTTTGTGATTAAGGAATCTT 18617
QY 1637 GACTACTTTTATATCTTAAAGGAATATTTCTTTTATATATCTTCAAAATTTAGAACTTAACCTT 1696
Db 18618 GACTACTTTTATATCTTAAAGGAATATTTCTTTTATATATCTTCAAAATTTAGAACTTAACCTT 18677
QY 1697 AAAAGTTTTCTTCTGTAATTTGTAACGGGTGATTTATTAACCTTAGATAACAGAT 1756
Db 18678 AAAAGTTTTCTTCTGTAATTTGTAACGGGTGATTTATTAACCTTAGATAACAGAT 18737
QY 1757 ACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGAAATTTCTATTAAAAATTTTAACTGTTG 1816
Db 18738 ACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGAAATTTCTATTAAAAATTTTAACTGTTG 18797
QY 1817 TATTCCTTTTCAATGGGTGATGTCAGGGTGATTAACACAGACATTCATGGAAGGCATGCAG 1876
Db 18798 TATTCCTTTTCAATGGGTGATGTCAGGGTGATTAACACAGACATTCATGGAAGGCATGCAG 18857
QY 1877 TTTGTCCAATGTGACAGTTTGTAAATAAAACCATACACATTTTATTTAAGATTAAAA 1936
Db 18858 TTTGTCCAATGTGACAGTTTGTAAATAAAACCATACACATTTTATTTAAGATTAAAA 18917
QY 1937 TCTAACTGGAAGTCAGCTTGGAAAAATGGAATTTCCAAAGTATGTTTGGTGAGTCACAGA 1996
Db 18918 TCTAACTGGAAGTCAGCTTGGAAAAATGGAATTTCCAAAGTATGTTTGGTGAGTCACAGA 18977
QY 1997 TATTAATAATAGAAATTTCTGATGAGAGTTTTCAGTTTAAATACCAAGTCCTTTAGAGTCT 2056
Db 18978 TATTAATAATAGAAATTTCTGATGAGAGTTTTCAGTTTAAATACCAAGTCCTTTAGAGTCT 19037
QY 2057 TAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAAATACGTAAACCTTATAGAAATTAAG 2116
Db 19038 TAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAAATACGTAAACCTTATAGAAATTAAG 19097
QY 2117 TTTTAAATAGGCAATTTATGCTGTGATATTTCTTACGGGAGAAAGAGGATTTGATTG 2176
Db 19098 TTTTAAATAGGCAATTTATGCTGTGATATTTCTTACGGGAGAAAGAGGATTTGATTG 19157
QY 2177 GAAAGCAGTTTGGGAGAAAGTCTGCTGAAATTTCCAGAAATTTAAATGATTGTTTACAT 2236
Db 19158 GAAAGCAGTTTGGGAGAAAGTCTGCTGAAATTTCCAGAAATTTAAATGATTGTTTACAT 19217
QY 2237 AAACCTTTTGTACTTCAG 2253
Db 19218 AAACCTTTTGTACTTCAG 19234

```

; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758DIV2
; CURRENT APPLICATION NUMBER: US/10/339,656
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: 10/109,854
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 21234
; TYPE: DNA
; ORGANISM: Homo sapien
; US-10-339-656-3

Query Match      48.6%; Score 1097; DB 4; Length 21234;
Best Local Similarity 100.0%; Pred. No. 5.4e-277;
Matches 1097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1157 AGGAATTATGCTTTGTCATGATGAAGAAACATGAGAACTGTTTGAACCTGGTTCGAAGAA 1216
DB 18138 AGGAATTTATGCTTTGTCATGATGAAGAAACATGAGAACTGTTTGAACCTGGTTCGAAGAA 18197

QY 1217 TGTTAGATATGATCCCACTCAAGAAATACCTTGGATGAAGCATGCGAGCATCCTTTCT 1276
DB 18198 TGTTAGATATGATCCCACTCAAGAAATACCTTGGATGAAGCATGCGAGCATCCTTTCT 18257

QY 1277 TTGACTTATTAAAAAGAAATGAAATGGAATCACTGCTGCTTACTATATATCTTCTAGA 1336
DB 18258 TTGACTTATTAAAAAGAAATGAAATGGAATCACTGCTGCTTACTATATATCTTCTAGA 18317

QY 1337 AGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTCCTAATATTTTGTAAACATTA 1396
DB 18318 AGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTCCTAATATTTTGTAAACATTA 18377

QY 1397 TTATTTTGTACAGTTAAAGTGAAATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1456
DB 18378 TTATTTTGTACAGTTAAAGTGAAATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 18437

QY 1457 TTAAGCAAGTATGCTTTGATAATGCAATAGAAAAATTAATAATTTCTTTTGTAA 1516
DB 18438 TTAAGCAAGTATGCTTTGATAATGCAATAGAAAAATTAATAATTTCTTTTGTAA 18497

QY 1517 ATTACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATTAATGTGATGATC 1576
DB 18498 ATTACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATTAATGTGATGATC 18557

QY 1577 TTGCTTTTGTACATGGAAGTCACTCTGAAGTGATTTTGTGAGTAAAGGAATCTT 1636
DB 18558 TTGCTTTTGTACATGGAAGTCACTCTGAAGTGATTTTGTGAGTAAAGGAATCTT 18617

QY 1637 GACTACTTTTATATCTTAAAGGAATATCTTTTATATATCTTCAAAATTTAGAACTTA 1696
DB 18618 GACTACTTTTATATCTTAAAGGAATATCTTTTATATATCTTCAAAATTTAGAACTTA 18677

QY 1697 AAAAGTTTCTCTGTAATTTGTTGAACGGGTGATTAATTAATCTAGATAAGCAGGT 1756
DB 18678 AAAAGTTTCTCTGTAATTTGTTGAACGGGTGATTAATTAATCTAGATAAGCAGGT 18737

QY 1757 ACTAGAAACCAAACTCAGAAATGTTTACTGTTAGAAATCTTATTAATTTTAAAGTTG 1816
DB 18738 ACTAGAAACCAAACTCAGAAATGTTTACTGTTAGAAATCTTATTAATTTTAAAGTTG 18797

QY 1817 TATTTCTTTTCAATTTGGGTGATGTCAGGGTGATTAACAGACATTTCAATGGAAGGCATG 1876
DB 18798 TATTTCTTTTCAATTTGGGTGATGTCAGGGTGATTAACAGACATTTCAATGGAAGGCATG 18857

QY 1877 TTGTCCATTTGTGACAGTTTGTGTTAATTAACCAATACATACACATTTTATTAAGATTA 1936
DB 18858 TTGTCCATTTGTGACAGTTTGTGTTAATTAACCAATACATACACATTTTATTAAGATTA 19917

```

```

QY 1937 TCTAACTGGAAAGTCAGCTTGGAAAAATGGACATTTCCAAGTATGTTTGGTGAGTCACAGA 1996
DB 18918 TCTAACTGGAAAGTCAGCTTGGAAAAATGGACATTTCCAAGTATGTTTGGTGAGTCACAGA 18977

QY 1997 TATAAAAATAGAAATTCCTGATGAGAGGTTTCAGTTTTTAATACCAAGTCTCTTAGAGTCT 2056
DB 18978 TATAAAAATAGAAATTCCTGATGAGAGGTTTCAGTTTTTAATACCAAGTCTCTTAGAGTCT 19037

QY 2057 TAACATTTGGCCAGCATCTGTTTATCAAAATGACATATAAATACGTAACCTTATAAGAAATTAAG 2116
DB 19038 TAACATTTGGCCAGCATCTGTTTATCAAAATGACATATAAATACGTAACCTTATAAGAAATTAAG 19097

QY 2117 TTTTATTAATTAGGCAATTTATGCTGCTGATAATTTCTTACGGGAGAAAGAGGATTTGATG 2176
DB 19098 TTTTATTAATTAGGCAATTTATGCTGCTGATAATTTCTTACGGGAGAAAGAGGATTTGATG 19157

QY 2177 GAAAGCAGTTTGGGAGAAAGAGTGTCTGCTGAAATTTCCAGAAATTAATGATGTTGATCAT 2236
DB 19158 GAAAGCAGTTTGGGAGAAAGAGTGTCTGCTGAAATTTCCAGAAATTAATGATGTTGATCAT 19217

QY 2237 AAACCTTTTGGACTTCAG 2253
DB 19218 AAACCTTTTGGACTTCAG 19234

```

RESULT 10

```

US-09-919-039-238
; Sequence 238, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE OF INVENTION: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 238
; LENGTH: 2254
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 420527.25
; NAME/KEY: unsure
; LOCATION: 231, 241
; OTHER INFORMATION: a, t, c, g, or other
US-09-919-039-238

```

```

Query Match      28.4%; Score 640.6; DB 4; Length 2254;
Best Local Similarity 72.1%; Pred. No. 1e-157;
Matches 880; Conservative 0; Mismatches 329; Indels 12; Gaps 3;

QY 327 ATATGAAATCTGTGACACACTTTGGTGAAGAGGCTTTTGGCAAAAGTTGTAGAGTGCAATGA 386
DB 983 AGATGAAATCTGTGATATCTTTAGGTGAAGAGCTTTTGGAAAAGTTGTGGAGTGCAATGA 1042

QY 387 TCATGGCATGGATGGCATGTCATGTAGCAGTGAATAATCGTAAAAAATGTAGCCCGTACCG 446
DB 1043 TCATAAAGCGGAGGTAGACATGTAGCAGTAAAAAATAGTAAAAAATGTGGATAGATACGT 1102

QY 447 TGAAGCAGCTCGTTCAGAAATCCAGATTTAGAGCAGTTAAATAGTACTGATCCCAATAG 506
DB 1103 TGAAGCTCTCGCTCAGAAATACAGTTCTGGAAATCTCTGAATACACAGACCCCAACAG 1162

QY 507 TGCTTCCGATGTGCCAGATGCTAGAAATGGTTTGATCATCATGCTGATGTTGTTGTTGTTGTT 566
DB 1163 TACTTTCGCTGTGTCAGATGTTGGAATGTTTGAGCATCATGTCACATTTGCTGATTTG 1222

```

```
QY 567 GTTGAACCTACTGGGACTTAGTACTACGATTTTCAATTAAGAAACAGCTTTCTGCCATT 626
Db 1223 TTTTGAACCTATTGGGACTTAGTACTACGATTTTCAATTAAGAAATGGTTTCTACCACT 1282
QY 627 TCAAAATGACCAATCAGGCAGATGGCGTATCAGATCTGCCAGTCAATAAATTTTTCACA 686
Db 1283 TCGACTGGATCATATCAGAAAGATGGCATATCAGATATGCAATCTGTGAATTTTGTGA 1342
QY 687 TCATTAATTAATTAACCCATACAGATCTGAAGCCTGAAATATTTTGTGTGAAGTCTGA 746
Db 1343 CAGTAATAAGTTGACTCACACAGACTTAAGCCTGAAACATCTTATTTGTGCACTGTA 1402
QY 747 CTATGTAGTCAATATAATTTCTAAATGAAAGCTGATGACCGCACACTGAAACACAGA 806
Db 1403 CTACACAGGCGGTATAATCCCAAAATAAAGCTGATGAACGACCTTAATAAATCCAGA 1462
QY 807 TATCAAGTTGTGACTTTGGAAGTGCACAGCTATGATGATGAATCAATCAAGTACTTTGGT 866
Db 1463 TATTAAGTTGTGACTTTGGAAGTGCACATATGATGACGAACATCAAGTACTTTGGT 1522
QY 867 GTCTACCGGCACTACAGAGTCCCGAGTCAATTTTGGCTTTAGTTGGTCTCAGCCCTTG 926
Db 1523 ATCTACAAGACATTATAGACACCTGAAAGTTATTTTAGCCCTAGGCTGCCAACCATG 1582
QY 927 TGATGTTTGGACATAGTTGCAATCTTATGAAATATTACCTTTGTTTACAGTCTTTCA 986
Db 1583 TGATGCTGGGACATAGGATGCAATCTTATTAATTAATCTTCTTGGGTTTACCGTATTTCC 1642
QY 987 GACTCATGATGTAAGAGCACCTCGCAATGATGGAACGAATATTAGGACCCATACACCA 1046
Db 1643 AACACAGATAGTAGAGGACATTTAGCAATGATGGAAGGATCTTTGGACCTCTACCAA 1702
QY 1047 ACACATGATTTACAGAAAAACAGAAACGCAAGTATTTTCAACCAATACAGCTAGATTGGGA 1106
Db 1703 ACATATGATACAGAAAAACAGGAACGTAATTTTCAACCAAGATGATTTAGACTGGGA 1762
QY 1107 TGAACACAGTCTGCTGTGATGATGTTAGGAGACGCTGCAAAACCGTTGAAGGAATTTAT 1166
Db 1763 TGAACACAGTCTGCTGCGGAGATATGTTTCAAGACGCTGTAAACCTCTGAAGGAATTTAT 1822
QY 1167 GCTTTGTGATGATGAAGAACATGAGAACTGTTTCAACCTGTTTGAAGAACTTTAGATA 1226
Db 1823 GCTTTCTCAAGATGTTGAACATGAGCGTCTCTTGAACCTCATTCAGAAAAATGTTGAGTA 1882
QY 1227 TGATCCAACTCAAGAAATTAACCTTGGATGAAGCATTTGAGCATCTCTTTTGTGACTTAT 1286
Db 1883 TGATCCAGCAAAAGAAATTAACCTCAGAGAACCTTTAAGCATCTCTTTTGTGACTTCT 1942
QY 1287 AAAAAAGAAATGAATGGGAATCAGTGTCTTACTATATATCTTCTCTAGAGAGATTAAT 1346
Db 1943 GAAAGAAA-----GTATATAGATCTGTAATTTGGACAGCTCTCTCGAAGAGA-TCTT 1992
QY 1347 TAAGACTGTGTCAGTC--AACTAAACATCTTAATATTTTGTAAACATTAATATTTTG 1404
Db 1993 ACAGACTGTATCAGTCTTAATTTTAAATTTTAAAGTTATTTTGTACAGCTTTGTAAATCT 2052
QY 1405 TACAGTTAAGTGAATATTTGATGTTTGTATCAATAGCATAAATACTTGTAAAGCAA 1464
Db 2053 TAACATTTTATTTGCGCATGTTTATTTTGTGTTGGTAAATTTGTTTCAATAGTACATAG 2112
QY 1465 GTATGGCTGTGATAGCAATAGAAAAAATAAATTAATTTTCTTTTGTAAATTAACAT 1524
Db 2113 CTAAGGTAATGAACATCTTTTTCAGTAATTTGAAAGTGAATTTATTCAGAAATAAATTTT 2172
QY 1525 TTTTAATACCTTTGAATAT 1545
Db 2173 GTGCTTATGAAGTTGATATGT 2193
```

RESULT 11

US-09-905-999-22

; Sequence 22, Application US/09905999

; Patent No. 6797513

```
; GENERAL INFORMATION:
; APPLICANT: ULLRICH, Axel
; APPLICANT: NAVLER, Oliver
; TITLE OF INVENTION: CLK PROTEIN KINASES AND RELATED PRODUCTS AND METHODS
; FILE REFERENCE: 038602/0431
; CURRENT APPLICATION NUMBER: US/09/905,999
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 09/127,248
; PRIOR FILING DATE: 1999-07-31
; PRIOR APPLICATION NUMBER: PCT/IB97/00946
; PRIOR FILING DATE: 1997-06-17
; PRIOR APPLICATION NUMBER: US 08/877,150
; PRIOR FILING DATE: 1997-06-17
; PRIOR APPLICATION NUMBER: US 60/034,286
; PRIOR FILING DATE: 1996-12-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: Mus musculus
; US-09-905-999-22
```

Query Match 21.3%; Score 481.2; DB 4; Length 1538;

Best Local Similarity 64.6%; Pred. No. 5.5e-116;

Matches 736; Conservative 0; Mismatches 398; Indels 6; Gaps 1;

```
QY 143 ATCGAATCCACTGCGAGTAAATCTTTCAGTCCGAGGAGGAGGAGTCTCTAAAGGAAAGC 202
Db 329 ATCGAGAGAAACAGCAGTTACCGAGGCCAGCGAGCGGAGGAGGAGGAGGAGGAGGAG 388
QY 203 GCAATAGACATCTGTTCAAGTCATCAGTCACTGTTGGAAGAGCCAGGAGGAGGAGGAGG 262
Db 389 GGAGAGGAGCGGAGCATTTAGCGCTCATCTTCAACAGCAGCGG-----GAGAGGCA 442
QY 263 GGAGTATAGAGGATGATGAGGAGGAGTCACTGATCTGTCAAAGTGGAGGAGGAGGAGG 322
Db 443 AGAGTGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 502
QY 323 CAAGATATGAATCGTGACACACTTTGGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 382
Db 503 AGCGATATGAATTTGTAAGCACCTTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 562
QY 383 TTGATCATGGCATGGAATGGCATGATGAGCAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAG 442
Db 563 TGGACCATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 622
QY 443 ACCGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 502
Db 623 ACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 682
QY 503 ATAGTGTCTTCCGATGTTCCAGATGCTAGAAATGTTTGGATCATCATGTCATGTTTGTGA 562
Db 683 ACAACAGAACTCTGTGTCAGATGTTTGTGCTGTTGATCTACCATGCCACATGTTGTA 742
QY 563 TTGTTTGAACCTACTGGGACTTAGTACTTACGATTTTCAATTAAGAGGAGGAGGAGGAGG 622
Db 743 TCTCTTTGAGCTTCTGGGCTTAGCAGCTTTCGATTTTCTCAAAGAGGAGGAGGAGGAGGAG 802
QY 623 CATTTCAAATTTGACCAATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 682
Db 803 CCTACCCCATCAACCAAGTGGGCGCATGAGGCTTCCAGCTCTGCGGAGGAGGAGGAGGAG 862
QY 683 TACATCATATAAATTAACCCATACAGATCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 742
Db 863 TCCATGATACAGTTGACATACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 922
QY 743 CTGACTATGATCAAAATATAATTTCTAAATGAAACGTCGATGAAACGAGGAGGAGGAGGAG 802
Db 923 CAGACTACGAATCACCTACAACTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 982
QY 803 CAGATATCAAAAGTTGTTGACTTTTGAAGTGAACGTCATGATGATGATGATGATGATGAT 862
```


	Query Match	19.1%	Score 430.8;	DB 4;	Length 1763;
	Best Local Similarity	63.5%;	Pred. No. 1e-102;		
	Matches 657;	Conservative	0;	Mismatches 377;	Indels 0; Gaps 0;
Qy	246	CCGAAGGAAAGATCCAGGAGTAGTAGAGGATGATAGGAGGGTCACCTGATCTGTCTCAAG	305		
Db	440	CAGTAAGCGCAGACCGGAGTGTGGAAGATGACAAAGAGGGTCACCTGTGTGCCGAT	499		

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2005, 01:13:38 ; Search time 1192.6 Seconds
(without alignments)
11198.198 Million cell updates/sec

Title: US-10-801-671-1_COPY_72_2327

Perfect score: 2256

Sequence: 1 gacactcaccagtcatta.....aaacttttgacttcagaaa 2256

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004as:*
13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2256	100.0	2354	6 AAD32038	Aad32038 Human kin
2	2215.8	98.2	2497	4 AAF89402	Aaf89402 Human cel
3	2048	90.8	2488	4 AAF44675	Aaf44675 Novel pro
4	2048	90.8	2488	12 ADI29373	Adi29373 Human MAR
5	2046.6	90.7	2446	6 AAX89852	Aax89852 Human pro
6	1632	72.3	1814	13 ADP24393	Adp24393 PRO polyp
7	1629.4	72.2	1792	12 ADI26142	Adi26142 Human cdn
8	1300.6	57.7	1323	3 AAC81772	Aac81772 Human bea
9	1286	57.0	1446	12 ADO07809	Ado07809 Human pol
10	1277.4	56.6	1881	12 ADI26140	Adi26140 Human cdn
11	1185	52.5	1222	3 AAC59283	Aac59283 Human sec
12	1178.2	52.2	1456	11 ADI31373	Adi31373 Human cdn
13	1118.8	49.6	1865	12 ADI26144	Adi26144 Human cdn
14	1097.2	48.6	1446	12 ADO07812	Ado07812 Mouse pol
15	1097	48.6	21234	6 AAD32039	Aad32039 Human kin
16	813.2	36.0	1455	12 ADO07810	Ado07810 Human pol
17	811.6	36.0	1834	12 ADH58708	Adh58708 Human CDC
18	811.6	36.0	1834	13 ADRI4000	Adri4000 Human NF-
19	811.6	36.0	1834	13 ADP24718	Adp24718 PRO polyp
20	768.8	34.1	1452	12 ADO07813	Ado07813 Mouse pol

21	768.4	34.1	906	12 ADO07814	Ado07814 Mouse pol
22	669.2	29.7	1785	10 ADG74696	Adg74696 Human kin
23	644.4	28.6	1743	12 ADH58715	Adh58715 Human CDC
24	643.2	28.5	1750	6 ABK84044	Abk84044 Human cdn
25	643.2	28.5	1750	10 ADL24750	Adl24750 Intestina
26	640.6	28.4	2254	12 ADE77073	Ade77073 Human cdn
27	640	28.4	2516	3 AAF21734	Aaf21734 Human bre
28	630.6	28.0	1643	10 ADL24751	Adl24751 Intestina
29	499.6	22.1	1500	10 ACA63031	ACA63031 Human clk
30	499.6	22.1	1973	6 ABL64388	Abi64388 Stomach c
31	499.6	22.1	1973	6 ABN95692	Abn95692 Gene #219
32	499.6	22.1	1973	12 ADQ15173	Adq15173 Human can
33	499.6	22.1	1973	12 ADQ83230	Adq83230 Human tum
34	499.6	22.1	1973	12 ADQ83910	Adq83910 Human tum
35	499.6	22.1	1973	12 ADQ84393	Adq84393 Human tum
36	499.6	22.1	1973	12 ADQ85141	Adq85141 Human tum
37	499.6	22.1	1973	12 ADQ86323	Adq86323 Human tum
38	499.6	22.1	2111	9 AAD57339	Aad57339 Human kin
39	464.4	20.6	1296	4 AAH46906	Aah46906 cDNA enco
40	464.4	20.6	1842	5 AAS87541	Aas87541 DNA enco
41	461.6	20.5	1026	12 ADO07807	Ado07807 Human pol
42	458.6	20.3	1885	12 ADJ62819	Adj62819 Human cdn
43	452.2	20.0	1473	12 ADO07811	Ado07811 Mouse pol
44	443	19.6	443	3 AAC81771	Aac81771 Human bea
45	441.4	19.6	2905	8 ACC46217	Acc46217 Human dit

ALIGNMENTS

RESULT 1
AAD32038
ID AAD32038 standard; cDNA; 2354 BP.

AC AAD32038;

DT 18-JUN-2002 (first entry)

DE Human kinase protein cDNA.

KW Human; kinase protein; enzyme; cytostatic; osteopathic; gene expression;
colon-moderately differentiated adenocarcinoma; chromosome mapping; Gene;
drug screening; therapeutic; gene therapy; tissue typing; chromosome 5;
bone osteosarcoma; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
FH 5'UTR 1..32
FT 5'UTR /tag= a
FT CDS 33..1370
FT /tag= b
FT /product= "Human kinase protein"
FT 3'UTR 1371..2354
FT /tag= c

WO200216567-A2.

PD 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US026389.

XX 24-AUG-2000; 2000US-0227470P.

PR 19-MAR-2001; 2001US-00810671.

XX (APPL-) APPLERA CORP.

PI Yan C, Ye J, Ketchum KA, Di Francesco V, Beasley EM;

DR WPI; 2002-269354/31.

DR P-PSDB; AAE20170.

PT New human kinase proteins and nucleic acids, useful in drug screening

assays, identifying modulators of kinase activity or treating disorders characterized by absence or unwanted expression of the protein.

Claim 4; Fig 1; 81pp; English.

The invention relates to isolated human kinase proteins and nucleic acids. The nucleic acid and peptide sequences can be used as models for the development of human therapeutic targets, aid in the identification of therapeutic proteins and serve as targets for the development of human therapeutic agents that modulate kinase activity in cells and tissues that express the kinase. The nucleic acids are useful as probes or primers, in constructing recombinant vectors, for expressing antigenic portions of the proteins, chromosome mapping, drug screening, testing an individual for a genotype, and for gene therapy in patients containing cells that are aberrant in kinase gene expression. The proteins may be used in drug screening assays, in the identification of compounds that modulate, stimulate or inhibit kinase activity, in pharmacogenomic analysis, in treating disorders characterised by an absence or unwanted expression of the protein (bone osteosarcoma, or colon-moderately differentiated adenocarcinoma), and in generating antibodies specific for the peptides. Such antibodies can be used to detect the protein in situ, in vitro, or in cell lysate or supernatant, to isolate and purify the proteins from host cells, pharmacogenomic analysis, tissue typing, and in inhibiting protein function. The present sequence is human protein cDNA. Human kinase protein gene is located on chromosome 5

Sequence 2354 BP; 783 A; 370 C; 472 G; 729 T; 0 U; 0 Other;

Query Match 100.0%; Score 2256; DB 6; Length 2354;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GACACTCATCCAGTCATTATTAGAACAGAGTCCCTTGAATGACGAGATTATCGGGAC	60
DB	72	GACACTCATCCAGTCATTATTAGAACAGAGTCCCTTGAATGACGAGATTATCGGGAC	131
QY	61	CGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTAT	120
DB	132	CGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTAT	191
QY	121	CACAGAGACATTGAAGCGGGTATCGAATCCACTGCAGTAATCTTCAGTCGGAGCAGG	180
DB	192	CACAGAGACATTGAAGCGGGTATCGAATCCACTGCAGTAATCTTCAGTCGGAGCAGG	251
QY	181	AGAAGCAGTCCCTTAAAGGAAGCGCAATAGACTGTTCAGAGTCATCAGTTCAGGTCGAAG	240
DB	252	AGAAGCAGTCCCTTAAAGGAAGCGCAATAGACTGTTCAGAGTCATCAGTTCAGGTCGAAG	311
QY	241	AGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATGAGGAGGGTCACCTGATCTGT	300
DB	312	AGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATGAGGAGGGTCACCTGATCTGT	371
QY	301	CAAAGTGAGAGCGTTCTAAGAGCAAGATAGAAATCGTTGGCACTTTGGGTGAAGGAGCC	360
DB	372	CAAAGTGAGAGCGTTCTAAGAGCAAGATAGAAATCGTTGGCACTTTGGGTGAAGGAGCC	431
QY	361	TTTGGCAAGTGTAGAGTGCATTCATCATGCAATGAGTGCATGCAATGAGTGCATGAGTGA	420
DB	432	TTTGGCAAGTGTAGAGTGCATTCATCATGCAATGAGTGCATGCAATGAGTGCATGAGTGA	491
QY	421	ATCGTAAAAAATGTAGGCGGTTTACCGTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGAG	480
DB	492	ATCGTAAAAAATGTAGGCGGTTTACCGTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGAG	551
QY	481	CACTTAAATAGTACTGATCCCAATAGTGTCTCCGATGTGCCAGTGCAGATGAGTGGTTT	540
DB	552	CACTTAAATAGTACTGATCCCAATAGTGTCTCCGATGTGCCAGTGCAGATGAGTGGTTT	611
QY	541	GATCATCATGGTCAATGTTTGTATGTTGTTGAACCTACTGGGACTTAGTACTTACGATTTTC	600
DB	612	GATCATCATGGTCAATGTTTGTATGTTGTTGAACCTACTGGGACTTAGTACTTACGATTTTC	671
QY	601	ATTAAGAAAAACAGCTTTCTGCCATTTCAAATTTGACCAATCAGGAGATGGCGTATCAG	660

DB	672	ATTAAGAAAAACAGCTTTCTGCCATTTCAAATTTGACCAATCAGGAGATGGCGTATCAG	731
QY	661	ATCTGCCAGTCAATAAATTTTACATCAATAAATAAATTAACCCATACAGATCTGAAGCCT	720
DB	732	ATCTGCCAGTCAATAAATTTTACATCAATAAATAAATTAACCCATACAGATCTGAAGCCT	791
QY	721	GAATAATATTTTGTGTGAAGTCTGAGCTATCTAGTCAAAATATAATTTCTAAATGAACGT	780
DB	792	GAATAATATTTTGTGTGAAGTCTGAGCTATCTAGTCAAAATATAATTTCTAAATGAACGT	851
QY	781	GATGAACGACACTGAAACACAGATATCAAAGTTGTGATCTTTGGAAGTGAACGATAT	840
DB	852	GATGAACGACACTGAAACACAGATATCAAAGTTGTGATCTTTGGAAGTGAACGATAT	911
QY	841	GATGATGAACATCAGCTACTTCTGCTGCTACCCGGCACTACAGAGCTCCCGAGGTCAAT	900
DB	912	GATGATGAACATCAGCTACTTCTGCTGCTACCCGGCACTACAGAGCTCCCGAGGTCAAT	971
QY	901	TTGGCTTTAGGTTGGTCTCAGCCTTTGATGATCTTTGGAGCATAGGTTGCAATCTTATTGAA	960
DB	972	TTGGCTTTAGGTTGGTCTCAGCCTTTGATGATCTTTGGAGCATAGGTTGCAATCTTATTGAA	1031
QY	961	TATTACCTTTGGTTTCAAGTCTTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAG	1020
DB	1032	TATTACCTTTGGTTTCAAGTCTTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAG	1091
QY	1021	GAAAGCAATATTAGGACCCATACACACATGATTTCAAGAAACAAAGAAACGCAAGTAT	1080
DB	1092	GAAAGCAATATTAGGACCCATACACACATGATTTCAAGAAACAAAGAAACGCAAGTAT	1151
QY	1081	TTTCACCATTAACAGTAGATTGGGATGAAACACAGTTCTGCTGGTAGATATGTTAGGAGA	1140
DB	1152	TTTCACCATTAACAGTAGATTGGGATGAAACACAGTTCTGCTGGTAGATATGTTAGGAGA	1211
QY	1141	CGCTGCAAAACCGTTGAAGGAATTTATGCTTTGTCATGATGAAGAACATGAGAAACTGTTT	1200
DB	1212	CGCTGCAAAACCGTTGAAGGAATTTATGCTTTGTCATGATGAAGAACATGAGAAACTGTTT	1271
QY	1201	GACCTGGTTCCGAAGAAATGTTAGAAATATGATCAACTCAAGAAATTTACCTGGATGAAGCA	1260
DB	1272	GACCTGGTTCCGAAGAAATGTTAGAAATATGATCAACTCAAGAAATTTACCTGGATGAAGCA	1331
QY	1261	TTGAGCATCTCTTTCTTTGACTTTATTAAGAAAGAAATGAAATGGGAATCAGTGGTCTTAC	1320
DB	1332	TTGAGCATCTCTTTCTTTGACTTTATTAAGAAAGAAATGAAATGGGAATCAGTGGTCTTAC	1391
QY	1321	TATATACTTCTAGAGAGATTTACTTTAAGACTGTGTCAGTCAACTAAACATCTCTAATAT	1380
DB	1392	TATATACTTCTAGAGAGATTTACTTTAAGACTGTGTCAGTCAACTAAACATCTCTAATAT	1451
QY	1381	TTTTGTAACCATTAATTTATTTTGTACAGTTAAAGTGTAAATATTTGATGTTTGTATCAA	1440
DB	1452	TTTTGTAACCATTAATTTATTTTGTACAGTTAAAGTGTAAATATTTGATGTTTGTATCAA	1511
QY	1441	TAGCATATTAACCTTTGTTAAGCAAGTATGGTCTTTGATATGCAATTTAGAAAAATTTAAAT	1500
DB	1512	TAGCATATTAACCTTTGTTAAGCAAGTATGGTCTTTGATATGCAATTTAGAAAAATTTAAAT	1571
QY	1501	AATTTTCTTTTGAATTTACATTTTAAATACCTTTGAAATATCTCTTTGTTGTCAGTG	1560
DB	1572	AATTTTCTTTTGAATTTACATTTTAAATACCTTTGAAATATCTCTTTGTTGTCAGTG	1631
QY	1561	ATAAATGTTGATTTCTTTGCTTTTGTACATGGAGTCACTCTGAAAGTGAATTTTTTTTG	1620
DB	1632	ATAAATGTTGATTTCTTTGCTTTTGTACATGGAGTCACTCTGAAAGTGAATTTTTTTTG	1691
QY	1621	AGTAAAGGAAATCTGACTTTTATATTTCTTAAAGAAATATTTCTTTTATATATCTTCAA	1680
DB	1692	AGTAAAGGAAATCTGACTTTTATATTTCTTAAAGAAATATTTCTTTTATATATCTTCAA	1751
QY	1681	TTTAGAACCTTAACCTTTAAAGCTTTTCTTCTGTAATCTGTAACGCGGTGATTATTATTA	1740

Db	1752	TTTGAAGACTTAACTTTTAAAAGTGTTCCTTCGTGTAATTTGTTGAACCGGTGATTTATTATTA	1811
Qy	1741	CTCTAGATAAGCAGGTACTAGAAACCAAACCTCAGAAAAATGTTTACTGTGTAGAAATCTCAT	1800
Db	1812	CTCTAGATNAGCAGGTACTAGAAACCAAACCTCAGAAAAATGTTTACTGTGTAGAAATCTCAT	1871
Qy	1801	TAAATTTTAAAGTGTGTATTCCTTTTTCATTTGGGGTGATGTCAGGGTGATTAACCGACATTC	1860
Db	1872	TAAATTTTAAAGTGTGTATTCCTTTTTCATTTGGGGTGATGTCAGGGTGATTAACCGACATTC	1931
Qy	1861	ATGGAAGGCATGCAGTTTGTCCATTCGTGACAGTTTCTTTTAAATAAAACCAATACACATTC	1920
Db	1932	ATGGAAGGCATGCAGTTTGTCCATTCGTGACAGTTTCTTTTAAATAAAACCAATACACATTC	1991
Qy	1921	TTATTTAAGATTAAATCTAACTGGAAGTCAGCTTGGAAAAATGGACATTTTCCAAAGTATG	1980
Db	1992	TTATTTAAGATTAAATCTAACTGGAAGTCAGCTTGGAAAAATGGACATTTTCCAAAGTATG	2051
Qy	1981	TTTGGTGAGTCACAGATATAAAAATAGAAATCTGTATGAGAGGTTTCAGTTTTTAAATPACC	2040
Db	2052	TTTGGTGAGTCACAGATATAAAAATAGAAATCTGTATGAGAGGTTTCAGTTTTTAAATPACC	2111
Qy	2041	AAGTCCTTAGGAGTCTTAAACATTTGGCCAGCATCTGTTTATCAAAATGACATAATACGTAA	2100
Db	2112	AAGTCCTTAGGAGTCTTAAACATTTGGCCAGCATCTGTTTATCAAAATGACATAATACGTAA	2171
Qy	2101	ACCTATAAGAAATTAAGTTTATTAATTAAGGCAATTTATGTCGTGATAAATCTTACGGGAG	2160
Db	2172	ACCTATAAGAAATTAAGTTTATTAATTAAGGCAATTTATGTCGTGATAAATCTTACGGGAG	2231
Qy	2161	AAAGAGGATTTGATTGGAAAGCAGTTTGGGAAGAAAGTCGTGCTGAAATTTTCCAGAAATTT	2220
Db	2232	AAAGAGGATTTGATTGGAAAGCAGTTTGGGAAGAAAGTCGTGCTGAAATTTTCCAGAAATTT	2291
Qy	2221	AATTGATTGGTTACATAAACCTTTTGTGACTTCAGAAA	2256
Db	2292	AATTGATTGGTTACATAAACCTTTTGTGACTTCAGAAA	2327

RESULT 2

AAF89402
ID AAF89402 standard; cDNA; 2497 BP.

XX

AC AAF89402; vv

DT 14-AUG-2001 (first entry)

XX Human cell cycle regulating protein 53 coding sequence.

XX Human; cell cycle regulating protein 53; cancer; blood disease; HIV;
KW immunological disease; inflammation; ss.
KW

X

OS Homo sapiens.
YY

PN WO200130833-

XX

PD
yy
03-MAY-2001.

PF 16-OCT-2000: 2000WO-CN000328.

[illegible]

PR 22-OCT-1999; 99CN-

XX
PA
(SHAN-) SHANGHAI

XX
PI Mao Y, Xie Y;
XX
XX WPI; 2001-300480/31.
DR P-PSDB; AAB69705.
DR
XX
PT New cell cycle-regulating protein 53 and its polynucleotide, applicable
in diagnosis and treatment of malignant tumor, hemopathy, human
PT immunodeficiency virus infection, immunological diseases and various
PT inflammation.

CC	activity include rheumatoid arthritis, atherosclerosis, autoimmune
CC	disorders, complications of organ transplantation, myocardial infarction,
CC	immune disorders, cardiomyopathies, strokes, renal failure, oxidative-
CC	stress related disorders, chronic inflammatory bowel disease, chronic
CC	inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis,
CC	psoriasis, rhinitis, autoimmunity, diabetes, cancers and reproductive
CC	disorders
xx	
sq	Sequence 2488 BP; 819 A; 400 C; 504 G; 764 T; 0 U; 1 Other;
	Query Match 90.8%; Score 2048; DB 4; Length 2488;
	Best Local Similarity 95.9%; Pred. No. 0;
	Matches 2148; Conservative 0; Mismatches 1; Indels 90; Gaps 1
Qy	14 GTCATTATTAGAACCAAGGTCCTTGAATGACGAGATTATCGGACCGGAGATACGTTG 73
Db	338 GTCATTATTAGAACCAAGGTCCTTGAATGACGAGATTATCGGACCGGAGATACGTTG 397
Qy	74 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTAGACATTATCACAGACATTG 133
Db	398 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTAGACATTATCACAGACATTG 457
Qy	134 AAGCGGGTATCGAATCCACTCGACGTAAATCTTCAGTCCGACGACGAGACGAGTCCTA 193
Db	458 AAGCGGGTATCGAATCCACTCGAGTAAATCTTCAGTCCGACGAGACGAGTCCTA 517
Qy	194 AAGGAAGCGCAATAGACACTGTTCAAGTCATCAGTCACGTTCCGAAGAGCCACCGAAGGA 253
Db	518 AAGGAAGCGCAATAGACACTGTTCAAGTCATCAGTCACGTTCCG----- 561
Qy	254 AAGATCCAGGAGTATAGAGSAGATGATGAGGAGGGTCACCTGATCTGTCAAAGTGGAGACG 313
Db	562 ----- 561
Qy	314 TTCTAGACCAAGATATGAATCGTGGACACTTTGGGTGAAGAGACCTTTGGCAAAGTTG 373
Db	562 -----NATGAAATCGTGGACACTTTGGGTGAAGAGACCTTTGGCAAAGTTG 607
Qy	374 TAGAGTGCAATTGATCATGCGATGGATGGCATGTCATGTAGCAGTCAAAATCGTAAAAATG 433
Db	608 TAGAGTGCAATTGATCATGCGCATGGATGGCATGTCATGTAGCAGTCAAAATCGTAAAAATG 667
Qy	434 TAGCGCGTTACCGTGAAGCAGCTCGTTCAGAAATCCAAATATTAGACACCTTAAATAGTA 493
Db	668 TAGCGCGTTACCGTGAAGCAGCTCGTTCAGAAATCCAAATATTAGACACCTTAAATAGTA 727
Qy	494 CTGATCCCAATAGTGTCTCCGATGTGTCCAGATGTCAGATGTTGTGATCATCATGGTC 553
Db	728 CTGATCCCAATAGTGTCTTCGATGTGTCCAGATGTCAGATGTTGTGATCATCATGGTC 787
Qy	554 ATGTTTGATTTGTTTGAACCTACTGGGACTTAGTACTCTACGATTTCCATTTAAAGAAACA 613
Db	788 ATGTTTGATTTGTTTGAACCTACTGGGACTTAGTACTCTACGATTTCCATTTAAAGAAACA 847
Qy	614 GCTTCTGCGCAATTTCAAATTTGACCAATCAGGCGAGATGGCGTATCAGATCTGCCAGTCAA 673
Db	848 GCTTCTGCGCAATTTCAAATTTGACCAATCAGGCGAGATGGCGTATCAGATCTGCCAGTCAA 907
Qy	674 TAAATTTTTTACATCATATAAATTAACCCATACAGATCTGAGCGCTGAAAAATATTTTGT 733
Db	908 TAAATTTTTTACATCATATAAATTTAAACCCATACAGATCTGAGCGCTGAAAAATATTTTGT 967
Qy	734 TTGTGAAGTCTGACATGTAGTCAAAATATAATTTCTTAAATGAAACGATGAACGCACAC 793
Db	968 TTGTGAAGTCTGACATGTAGTCAAAATATAATTTCTTAAATGAAACGATGAACGCACAC 102
Qy	794 TGAATAACACAGATATCAAAGTTGTGACTTTTGAAGTGCAACCGTATGATGATGAACATC 853
Db	1028 TGAATAACACAGATATCAAAGTTGTGACTTTTGAAGTGCAACCGTATGATGATGAACATC 108
Qy	854 ACAGTACTTTGGTGTCTACCGGCACTACAGAGCTCCGAGGTCATTTGGCTTTAGGTT 913
Db	1088 ACAGTACTTTGGTGTCTACCGGCACTACAGAGCTCCGAGGTCATTTGGCTTTAGGTT 114

1208	TCACAGTCTTTTCAGACCTCATGATAGTAAAGAGCAGCTGGCNAATGATGGAAAGAAATATTAG	126	
Qy	1034	GACCCATACCAACAACATGATTCAGAAAAACAAGAAAAACGCAAGTATTTTCCACCATAACC	1093
Db	1268	GACCCATACCAACAACAATGATTCAGAAAAACAAGAAAAACGCAAGTATTTTCCACCATAACC	1327
Qy	1094	AGCTAGATTTGGATGAACACAGTTCTGCTGGTAGATATGTTAGGAGACGCTGCMAACCGT	1153
Db	1328	AGCTAGATTTGGATGAACACAGTTCTGCTGGTAGATATGTTAGGAGACGCTGCMAACCGT	1387
Qy	1154	TGAAGGAATTTATGCTTTGTCATGATGAAGAAACATGAGAAACCTGTTTGACTCGTGTCCAA	1213
Db	1388	TGAAGGAATTTATGCTTTGTCATGATGAAGAAACATGAGAAACCTGTTTGACTCGTGTCCAA	1447
Qy	1214	GAATGTTAGAAATATGATCCAACTCAAAAGAAATACCTTTGGATGAAGCAATTCGACGATCCTT	1273
Db	1448	GAATGTTAGAAATATGATCCAACTCAAAAGAAATACCTTTGGATGAAGCAATTCGACGATCCTT	1507
Qy	1274	TCCTTGACTTATTTAAAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATATCTTCTCT	1333
Db	1508	TCCTTGACTTATTTAAAAAGAAATGAAATGGGAATCAGTGGTCTTACTATATATCTTCTCT	1567
Qy	1334	AGAAGAGATTACTTTAAGACTGTGTCAGTCAACTAAACATTCCTAATAATTTTTGTAAACATT	1393
Db	1568	AGAAGAGATTACTTTAAGACTGTGTCAGTCAACTAAACATTCCTAATAATTTTTGTAAACATT	1627
Qy	1394	AAATTTATTTTGTAACAGTTAAGTGTAAATAATTTGTAATGTTTGTATCAATAGCAATAATTAAC	1453
Db	1628	AAATTTATTTTGTAACAGTTAAGTGTAAATAATTTGTAATGTTTGTATCAATAGCAATAATTAAC	1687
Qy	1454	TTGTTAAGCAAGTATGGTCTTGATAATGCAATAGAAAAATTTAAATTTAAATTTTCTTTTTT	1513
Db	1688	TTGTTAAGCAAGTATGGTCTTGATAATGCAATAGAAAAATTTAAATTTAAATTTTCTTTTTT	1747
Qy	1514	GAATTTACCATTATTAATACCTTTGAAATATCCTTTGTGTCAGTGATATAATGTGATG	1573
Db	1748	GAATTTACCATTATTAATACCTTTGAAATATCCTTTGTGTCAGTGATATAATGTGATG	1807
Qy	1574	ATCTTGCTCTTTGTACATGGAGGTCACTCTGAAAGTGAATTTTTTTTTCAGTAAAGGAAAT	1633
Db	1808	ATCTTGCTCTTTGTACATGGAGGTCACTCTGAAAGTGAATTTTTTTTTCAGTAAAGGAAAT	1867
Qy	1634	CTTGACTACTTTTATTTCTTTAAAGGAATATCTTTTATATATCTTCAAAATTTAGAACTTAAC	1693
Db	1868	CTTGACTACTTTTATTTCTTTAAAGGAATATCTTTTATATATCTTCAAAATTTAGAACTTAAC	1927
Qy	1694	TTTTAAAGTTTTTCTCTGTAAATTTGTAACGGGTGATTTATTAATCTCTAGATAGCA	1753
Db	1928	TTTTAAAGTTTTTCTCTGTAAATTTGTAACGGGTGATTTATTAATCTCTAGATAGCA	1987
Qy	1754	GGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTGTAGAAATCTATAATTTTAAAGT	1813
Db	1988	GGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTGTAGAAATCTATAATTTTAAAGT	2047
Qy	1814	TTGTATTCTTTTTTCATTTGGGTGATGTGAGGTGATAACCGACATTCATCGAAAGGCATG	1873
Db	2048	TTGTATTCTTTTTTCATTTGGGTGATGTGAGGTGATAACCGACATTCATCGAAAGGCATG	2107
Qy	1874	CAGTTTGTCCATTTGTGACAGTTTGTGTTAAATAAAACCAATACACACTTTTATTTAAGATTA	1933
Db	2108	CAGTTTGTCCATTTGTGACAGTTTGTGTTAAATAAAACCAATACACACTTTTATTTAAGATTA	2167
Qy	1934	AAATCTAACTGGAAGTCAGCTTCGAAAAATGGACATTTTCCAAGTATGTTTGGGTGAGTCAC	1993
Db	2168	AAATCTAACTGGAAGTCAGCTTCGAAAAATGGACATTTTCCAAGTATGTTTGGGTGAGTCAC	2227
Qy	1994	AGATATAAAATAGAAATTTCTGATGAGAGGTTTTCAGTTTTTAAATACCAAGTCTCTTAGGAG	2053
Db	2228	AGATATAAAATAGAAATTTCTGATGAGAGGTTTTCAGTTTTTAAATACCAAGTCTCTTAGGAG	2287
Qy	2054	TCTTAAACATTTGGCCAGCATCTGTTTTATCAAAATGACATAAATACGTAACCTATAAGAAATT	2113
Db	2288	TCTTAAACATTTGGCCAGCATCTGTTTTATCAAAATGACATAAATACGTAACCTATAAGAAATT	2347

Qy	2114	AAGTTTAAATTAAGGCAATTATATGCTCTGTGATTAATCTTACGGGAGAAAGAGGATTGGA	2173
Db	2348	AAGTTTAAATTAAGGCAATTATATGCTCTGTGATTAATCTTACGGGAGAAAGAGGATTGGA	2407
Qy	2174	TTGGAAGACGAGTTTGGGAAGAAAGTCTGCTGAGAAATTTCCAGAAATTTAAATTGATTGGTTA	2233
Db	2408	TTGGAAGACGAGTTTGGGAAGAAAGTCTGCTGAGAAATTTCCAGAAATTTAAATTGATTGGTTA	2467
Qy	2234	CATAAACTTTTGTGACTTCA	2252
Db	2468	CATAAACTTTTGTGACTTCA	2486
RESULT 5			
AAx89852			
ID	AAx89852 standard; cDNA; 2446 BP.		
XX			
AC	AAx89852;		
XX			
DT	08-OCT-1999 (first entry)		
XX			
DE	Human protein kinase (HPKM)-3 encoding nucleotide (clone ID 339963).		
XX			
KW	Human protein kinase molecule; HPKM; human; protein kinase;		
KW	phosphate group; cancer; immune disorder; ss.		
XX			
OS	Homo sapiens.		
XX			
FH	Key Location/Qualifiers		
FT	CDS 128..1483		
FT	/*tag= a		
FT	/product= "HPKM-3"		
XX			
PN	W09938981-A2.		
XX			
PD	05-AUG-1999.		
XX			
PF	12-JAN-1999; 99WO-US0000661.		
XX			
PR	30-JAN-1998; 98US-00016000.		
XX			
PA	(INCY-) INCYTE PHARM INC.		
XX			
PI	Bandman O, Hillman JL, Lal P, Akerblom IE, Shah P, Corley NC;		
PI	Guegler KJ;		
XX			
DR	WPI; 1999-479190/40.		
DR	P-PSDB; AAY27054.		
XX			
PT	New human protein kinase molecules useful for treating or preventing		
PT	cancer or an immune disorder.		
XX			
PS	Claim 8; Page 74-75; 77pp; English.		
XX			
CC	The invention provides human protein kinase molecules (HPKM) (AAY27052-		
CC	57) and nucleic acid sequences (AAx89850-55) encoding the HPKM		
CC	polypeptides respectively. The HPKM polypeptides can be produced		
CC	recombinantly by standard recombinant methodology. Protein kinases add		
CC	phosphate groups to proteins. HPKM polypeptide or an HPKM antagonist is		
CC	used to treat or prevent cancer or an immune disorder. The present		
CC	sequence represents the nucleotide sequence of HPKM-3		
XX			
SQ	Sequence 2446 BP; 788 A; 391 C; 516 G; 751 T; 0 U; 0 Other;		

Query Match	90.7%	Score	2046.6	DB 2	Length	2446
Best Local Similarity	95.8%	Pred. No.	0			
Matches	2149	Conservative	0	Mismatches	4	Indels
					90	Gaps
QY	14	GTCAATTATTAGAACCAAGTCTCTTGAATGAGCGAGATTATCGGACCGGAGATACGTTG	73			
DB	288	GTCAATTATTAGAACCAAGTCTCTTGAATGAGCGAGATTATCGGACCGGAGATACGTTG	347			

QY 74 ACGAATACAGGAATGACTCTGTGAAGGATATGTTCTTAGACATTTATCACAGACATTG 133
Db 348 ACGAATA CAGGAAATGACTACTGTGAAGGATATGTTCTTAGACATTTATCACAGACATTG 407
QY 134 AAAGCGGTATCGAATCCACTGCAGTAAATCTTCACTCCGACGAGGAGAGCAGTCCCTA 193
Db 408 AAAGCGGTATCGAATCCACTGCAGTAAATCTTCACTCCGACGAGGAGAGCAGTCCCTA 467
QY 194 AAAGGAAGCGCAATAGACACTGTTTCAAGTCATCAGTCAGTCTTCAAGAGGCCACCGAAGGA 253
Db 468 AAAGGAAGCGCAATAGACACTGTTTCAAGTCATCAGTCAGTCTTCAAGAGGCCACCGAAGGA 515
QY 254 AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTCGATCTGTCAAAAGTGAGAGC 313
Db 516 ----- 515
QY 314 TTCTAAGACGAAGATATGAATTCGTGGACACTTTGGGTGAAGGAGCCCTTTGGCAAGTTG 373
Db 516 -----AATCCGTGGACACTTTGGGTGAAGGAGCCCTTTGGCAAGTTG 557
QY 374 TAGAGTGCAATTGATCATCGCATGATGCGATGCGATGAGTGAAGTCAAAATCGTAAAGAAATG 433
Db 558 TAGAGTGCAATTGATCATCGCATGATGCGATGCGATGAGTGAAGTCAAAATCGTAAAGAAATG 617
QY 434 TAGGCGGTTACCGTGAAGCAGCTCGTTCCAGAAATCCAAAGTATTAGAGCACTTAAATAGTA 493
Db 618 TAGGCGGTTACCGTGAAGCAGCTCGTTCCAGAAATCCAAAGTATTAGAGCACTTAAATAGTA 677
QY 494 CTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGCTAGATGGTTTGGATCATCATGCTC 553
Db 678 CTGATCCCAATAGTGTCTTCCGATGTGTCCAGATGCTAGATGGTTTGGATCATCATGCTC 737
QY 554 ATGTTTGTATTGTTTGAACCTACCTGGGACTTAGTACTTACGATTTCAATTAAGAAAAACA 613
Db 738 ATGTTTGTATTGTTTGAACCTACCTGGGACTTAGTACTTACGATTTCAATTAAGAAAAACA 797
QY 614 GCTTTCTGCCATTTCAAAATGACCAATCAGCAGATGCGGTATCAGATCTGCCAGTCAA 673
Db 798 GCTTTCTGCCATTTCAAAATGACCAATCAGCAGATGCGGTATCAGATCTGCCAGTCAA 857
QY 674 TAAATTTTACATCATTAATAATTAACCCATACAGATCTGAGGCTCGAAATATTTTGT 733
Db 858 TAAATTTTACATCATTAATAATTAACCCATACAGATCTGAGGCTCGAAATATTTTGT 917
QY 734 TTGTGAAGTCTGACTATGTAGTCAAAATATAATCTTAAATGAAACGCTGATGAACCCAC 793
Db 918 TTGTGAAGTCTGACTATGTAGTCAAAATATAATCTTAAATGAAACGCTGATGAACCCAC 977
QY 794 TGAATAACACAGATATCAAAAGTTGTGACTTTGGAAAGTGCAACGATGATGATGAACATC 853
Db 978 TGAATAACACAGATATCAAAAGTTGTGACTTTGGAAAGTGCAACGATGATGATGAACATC 1037
QY 854 ACAGTACTTTGGTGTCTACCCGGGACTACAGAGCTCCCGAGTCAATTTTGGCTTTAGGTT 913
Db 1038 ACAGTACTTTGGTGTCTACCCGGGACTACAGAGCTCCCGAGTCAATTTTGGCTTTAGGTT 1097
QY 914 GGTCTCAGCCTTTGATGTTTGGAGCATAGGTTGCAATCTTATGAAATATTACCTTGGTT 973
Db 1098 GGTCTCAGCCTTTGATGTTTGGAGCATAGGTTGCAATCTTATGAAATATTACCTTGGTT 1157
QY 974 TCACAGTCTTTTCAGACTCATGATAGTAAAGAGCACCTGGCAATGATGGAACGAATATTAG 1033
Db 1158 TCACAGTCTTTTCAGACTCATGATAGTAAAGAGCACCTGGCAATGATGGAACGAATATTAG 1217
QY 1034 GACCCATACCAACACATGATTTCAAGAAACGAAGAAACGCAAGTATTTTCAACCAATACC 1093
Db 1218 GACCCATACCAACACATGATTTCAAGAAACGAAGAAACGCAAGTATTTTCAACCAATACC 1277
QY 1094 AGCTAGATTGGATGAACACAGTCTCTGCTGGTAGATATGTTAGGAGACGCTGCAACCCGT 1153
Db 1278 AGCTAGATTGGATGAACACAGTCTCTGCTGGTAGATATGTTAGGAGACGCTGCAACCCGT 1337
QY 1154 TGAAGGAATTTATGCTTTGTCATGATGAAGAAACATGAGAGAAACATGTTTGCCTGGTTCGAA 1213

Db 1338 TGAAGGAATTTATGCTTTGTCTATGATGAAGAAACATGAGAAACCTGTTGACCTGGTTCGAA 1397
QY 1214 GAATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGACATTTGCAGCATCCTT 1273
Db 1398 GAATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGACATTTGCAGCATCCTT 1457
QY 1274 TCTTTGACTTATTAATAAGAAATGAATGGAATCAGTGGTCTTACTATATATCTTCT 1333
Db 1458 TCTTTGACTTATTAATAAGAAATGAATGGAATCAGTGGTCTTACTATATATCTTCT 1517
QY 1334 AGAAGAGATTAATTAAGACTGTGTCAGTCAACTAAACATTTCTTAATATTTTGTAAACATTT 1393
Db 1518 AGAAGAGATTAATTAAGACTGTGTCAGTCAACTAAACATTTCTTAATATTTTGTAAACATTT 1577
QY 1394 AAATTTATTTGCTACAGTTAACTGTAATATTTGTAATGTTTGTATCAATAGCATATTAATAC 1453
Db 1578 AAATTTATTTGCTACAGTTAACTGTAATATTTGTAATGTTTGTATCAATAGCATATTAATAC 1637
QY 1454 TTGTTAAGCAAGTATGGTCTTTGATAATGCAATTAAGAAATTAATAATTTTCTTTT 1513
Db 1638 TTGTTAAGCAAGTATGGTCTTTGATAATGCAATTAAGAAATTAATAATTTTCTTTT 1697
QY 1514 GAAATTAACATTTTAAATACCTTTGAAATATCTTTGTGTCAGTGAATAATGTAATG 1573
Db 1698 GAAATTAACATTTTAAATACCTTTGAAATATCTTTGTGTCAGTGAATAATGTAATG 1757
QY 1574 ATCTTGCTTTGTACATGAGGTCACCTCTGAAAGTGAATTTTGTAGTAAAGAGAAAT 1633
Db 1758 ATCTTGCTTTGTACATGAGGTCACCTCTGAAAGTGAATTTTGTAGTAAAGAGAAAT 1817
QY 1634 CTGACTACTTTTATATCTTTAAAGCAATATCTTTATATATCTTCAAAATTTAGAACTTAAC 1877
Db 1694 TTTAAAGTTTTCTTCTGTAATTTGTTGAAACGGGTGATTTATTAATCTCTAGATAAGCA 1753
QY 1878 TTTAAAGTTTTCTTCTGTAATTTGTTGAAACGGGTGATTTATTAATCTCTAGATAAGCA 1937
Db 1754 GGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGAAATCTTATTAATTTTAAAGTG 1813
QY 1938 GGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGAAATCTTATTAATTTTAAAGTG 1997
Db 1814 TTGTAATCTTTTCAATTTGGGTGATGTCAAGGTGATTAACACATTCATGGAAGAGCATG 1873
QY 1998 TTGTAATCTTTTCAATTTGGGTGATGTCAAGGTGATTAACACATTCATGGAAGAGCATG 2057
Db 1874 CAGTTTGTCCATTTGTGACAGTTTGTTTAAATAAACCCACATACACATTTTATTAAGATTA 1933
QY 2058 CAGTTTGTCCATTTGTGACAGTTTGTTTAAATAAACCCACATACACATTTTATTAAGATTA 2117
Db 1934 AAATCTAACTGGAAGTCAGCTTGGAAAAATGGAACATTTTCCAAGTATGTTTGGTGAAGTCAC 1993
QY 2118 AAATCTAACTGGAAGTCAGCTTGGAAAAATGGAACATTTTCCAAGTATGTTTGGTGAAGTCAC 2177
Db 1994 AGATAAAAAATAGAAATTTCTGATGAGAGTTTCAAGTTTAAATACCAAGTCCTTAGGAG 2053
QY 2178 AGATAAAAAATAGAAATTTCTGATGAGAGTTTCAAGTTTAAATACCAAGTCCTTAGGAG 2237
Db 2054 TCTTAACATTTGGCCAGCATCTGTTTATCAAAATGATCAATTAACCTGTAACCTATTAAGAAAT 2113
QY 2238 TCTTAACATTTGGCCAGCATCTGTTTATCAAAATGATCAATTAACCTGTAACCTATTAAGAAAT 2297
Db 2114 AAGTTTATTAATTAAGCAATTTATGCTGTGATTAATCTTACGGGAGAAAGAGGATTTGA 2173
QY 2298 AAGTTTATTAATTAAGCAATTTATGCTGTGATTAATCTTACGGGAGAAAGAGGATTTGA 2357
Db 2174 TTGGAAGACAGTTTGGGAGAAAGTGTCTGTGAAATTTCCAGAAATTTAATTAATTTGGTTA 2233
QY 2358 TTGGAAGACAGTTTGGGAGAAAGTGTCTGTGAAATTTCCAGAAATTTAATTAATTTGGTTA 2417
Db 2234 CATAAACCTTTTGTGACTTCAGAAA 2256
QY |||||

Db 1181 GAACCATACCAACACATGATTCAGAAAAACGAAAGTATTTTCCACATAACC 1240
 Qy 1094 AGCTAGATTGGGATGAACACAGTTCTGCTGTGTAGATATGTTAGGAGCGCTGCAACCGT 1153
 Db 1241 AGCTAGATTGGGATGAACACAGTTCTGCTGTGTAGATATGTTAGGAGCGCTGCAACCGT 1300
 Qy 1154 TGAAGGAATTTATGCTTTGTGTCATGATGAAGAACATGAGAAACTGTTTGACCTGGTTGCAA 1213
 Db 1301 TGAAGGAATTTATGCTTTGTGTCATGATGAAGAACATGAGAAACTGTTTGACCTGGTTGCAA 1360
 Qy 1214 GAATGTTAGATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCAATTCGAGCATCCTT 1273
 Db 1361 GAATGTTAGATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCAATTCGAGCATCCTT 1420
 Qy 1274 TCTTTGACTTATTAATAAAGAAATGAATGGGATCAGTGTCTTACTATATCTTCTCT 1333
 Db 1421 TCTTTGACTTATTAATAAAGAAATGAATGGGATCAGTGTCTTACTATATCTTCTCT 1480
 Qy 1334 AGAAGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATT 1393
 Db 1481 AGAAGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATT 1540
 Qy 1394 AAATATTTTGTACAGTTAAGTGTAAATATTGTATGTTTGTATCAATAGCATATTAAC 1453
 Db 1541 AAATATTTTGTACAGTTAAGTGTAAATATTGTATGTTTGTATCAATAGCATATTAAC 1600
 Qy 1454 TTGTTAGCAAGTATGCTTGTATATGATCAATAGCAATTAATAATTTTCTTTT 1513
 Db 1601 TTGTTAGCAAGTATGCTTGTATATGATCAATAGCAATTAATAATTTTCTTTT 1660
 Qy 1514 GAAATACATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGAATAATGATGATG 1573
 Db 1661 GAAATACATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGAATAATGATGATG 1720
 Qy 1574 ATCTTGCCCTTTTGTCATGAGGTCACCTCTGAAGTGAATTTTGTGATGAAGGAAT 1633
 Db 1721 ATCTTGCCCTTTTGTCATGAGGTCACCTCTGAAGTGAATTTTGTGATGAAGGAAT 1780
 Qy 1634 CTGACTACTTATATTTCTTAAGGAATA 1662
 Db 1781 CTGACTACTTATTAATAAAGAAAAA 1809

RESULT 7
 ADI26142
 ID ADI26142 standard; cDNA; 1792 BP.
 AC ADI26142;
 XX
 DT 22-APR-2004 (first entry)
 DE Human cDNA encoding protein that promotes STAT6 activation #54.
 XX
 KW ss; gene; human; signal transducer and activator of transcription 6;
 KW STAT6; immunogen; STAT6 activation; allergy; inflammation;
 KW autoimmune disease; diabetes; hyperlipidaemia; infection; cancer;
 KW Th1 hyperactive disease; rheumatoid arthritis; osteoarthritis;
 KW systemic lupus erythematosus; sepsis; asthma; allergic rhinitis;
 KW ischaemic heart disease; subarachnoid haemorrhage; viral hepatitis; AIDS.
 OS Homo sapiens.
 XX
 PN WO2003104277-A2.
 XX
 PD 18-DEC-2003.
 XX
 PF 05-JUN-2003; 2003WO-JP007123.
 XX
 PR 05-JUN-2002; 2002JP-00164257.
 PR 06-JUN-2002; 2002US-0385912P.
 PR 26-DEC-2002; 2002JP-00377326.
 PR 27-DEC-2002; 2002US-0436467P.

PR 15-MAY-2003; 2003JP-00137505.
 PR 16-MAY-2003; 2003US-0470836P.
 XX (ASAH) ASAH KASEI KK.
 PA Sugahara T, Matsuda A, Honda G, Muramatsu S, Ishizawa K;
 FI WPI; 2004-122214/12.
 DR P-PSDB; ADI26143.
 DR
 XX
 PT New signal transducer and activator of transcription 6 activation
 PT promoting purified protein, for diagnosing and treating disease
 PT associated with activation/inhibition of transcription factor e.g.
 PT diabetes and cancer.
 XX
 PS Claim 4; SEQ ID NO 107; 1368bp; English.
 XX
 CC The invention relates to a purified protein promoting signal transducer
 CC and activator of transcription 6 activation (STAT6). The protein is
 CC useful for the producing an antibody, which involves administering the
 CC protein or its epitope-bearing fragments to a non-human animal as an
 CC antigen. The nucleic acid is useful for diagnosing a disease or
 CC susceptibility to a disease related to expression or activity of the
 CC protein. A transformant expressing the protein is useful for screening
 CC compounds which inhibit or promote STAT6 activation. A transformant
 CC expressing the protein is useful for producing a pharmaceutical
 CC composition. Compositions, antibodies and antisense molecules are useful
 CC for the treating a disease associated with STAT6 activation such as
 CC allergic diseases, inflammation, autoimmune diseases, diabetes,
 CC hyperlipidaemia, infectious disease and cancers. Compositions are useful
 CC for treating disease associated with STAT6 activation and/or prevention
 CC of Th1 hyperactive diseases. Compositions are also useful in rheumatoid
 CC arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma,
 CC allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage,
 CC viral hepatitis and AIDS. The protein has efficient promoting STAT6
 CC activity. The protein or nucleic acid is effectively useful for screening
 CC compounds for treating and preventing disease associated with excessive
 CC activation or inhibition of STAT6. The present sequence represents a
 CC human cDNA encoding a protein which promotes STAT6 activation.
 XX
 SQ Sequence 1792 BP; 587 A; 299 C; 385 G; 520 T; 0 U; 1 Other;

Query Match 72.2%; Score 1629.4; DB 12; Length 1792;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1630; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 14 GTCATTATTAGAACCAAGGTCCTTGAATGACGAGATTATCGGGACCGGAGATACGTTG 73
 Db 161 GTCATTATTAGAACCAAGGTCCTTGAATGACGAGATTATCGGGACCGGAGATACGTTG 220
 Qy 74 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTATCACAGAGACATTG 133
 Db 221 ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTTAGACATTATCACAGAGACATTG 280
 Qy 134 AAAGCGGTATCGAATCCACTGCAGTAAATCTTCAGTCCGACGAGGAGAGACGTCCTA 193
 Db 281 AAAGCGGTATCGAATCCACTGCAGTAAATCTTCAGTCCGACGAGGAGAGACGTCCTA 340
 Qy 194 AAAGGAAGCGCAATAGACACATGTTCAAGTCTCATCAGTCAAGTTCGGAAGCCACCGAAGGA 253
 Db 341 AAAGGAAGCGCAATAGACACATGTTTCAAGTCTCATCAGTCAAGTTCGGAAGCCACCGAAGGA 400
 Qy 254 AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACTGATCTGTCAAGTGTGAGACG 313
 Db 401 AAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACTGATCTGTCAAGTGTGAGACG 460
 Qy 314 TTCTAAGAGCAAGATATGAAATCGGGACACTTTTGGGTGAGAGCCCTTTCGCAAGTTG 373
 Db 461 TTCTAAGAGCAAGATATGAAATCGGGACACTTTTGGGTGAGAGCCCTTTCGCAAGTTG 520
 Qy 374 TAGAGTGCATTGATCATGCGCATGGAATGCGCATGTCATGATAGCAGTGAATAATCGTAAAAAATG 433
 Db 521 TAGAGTGCATTGATCATGCGCATGGAATGCGCATGTCATGATAGCAGTGAATAATCGTAAAAAATG 580

```
QY 434 TAGGCGTTACCGTGAAGCAGCTGCTTCAGAAATCCAAAGTATTAGAGCACTTTAAATAGTA 493
Db 581 TAGGCGTTACCGTGAAGCAGCTGCTTCAGAAATCCAAAGTATTAGAGCACTTTAAATAGTA 640
QY 494 CTGATCCCAATAGTCTCTCCGATGCTGTCAGATGCTAGAAATGGTTTCATCATCATGTC 553
Db 641 CTGATCCCAATAGTCTCTCCGATGCTGTCAGATGCTAGAAATGGTTTCATCATCATGTC 700
QY 554 ATGTTTGTATTGTTTGAACCTACTGGGACTTAGTACTTACGATTTCAATTAAGAAACA 613
Db 701 ATGTTTGTATTGTTTGAACCTACTGGGACTTAGTACTTACGATTTCAATTAAGAAACA 760
QY 614 GCTTCTGCCATTTCAAATTTGACCAATCAGCAGATGGCGTATCAGATCTGCCAGTCAA 673
Db 761 GCTTCTGCCATTTCAAATTTGACCAATCAGCAGATGGCGTATCAGATCTGCCAGTCAA 820
QY 674 TAAATTTTTCATCATATAATTAATTAACCCATACAGATCTGAAGCTGAAATATTTTGT 733
Db 821 TAAATTTTTCATCATATAATTAATTAACCCATACAGATCTGAAGCTGAAATATTTTGT 880
QY 734 TTGGAAGTCTGACTATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 793
Db 881 TTGGAAGTCTGACTATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 940
QY 794 TGAATAAACACAGATATCAAAAGTTGTGACTTTTGAAGTGCACCGTATGATGATGAATC 853
Db 941 TGAATAAACACAGATATCAAAAGTTGTGACTTTTGAAGTGCACCGTATGATGATGAATC 1000
QY 854 ACAGTACTTTGGTGTCTACCGGCACTACAGAGTCCCGAGGTCAATTTTGGCTTTAGTT 913
Db 1001 ACAGTACTTTGGTGTCTACCGGCACTACAGAGTCCCGAGGTCAATTTTGGCTTTAGTT 1060
QY 914 GGTCTCAGCTTTGATGTTTGGAGCATAGTTGCAATTTTATTGAATATTACCTTTGTT 973
Db 1061 GGTCTCAGCTTTGATGTTTGGAGCATAGTTGCAATTTTATTGAATATTACCTTTGTT 1120
QY 974 TCACAGTCTTTCAGACTCATAGTATTAAGAGCACTCGGCAATGATGGAACGAATATTAG 1033
Db 1121 TCACAGTCTTTCAGACTCATAGTATTAAGAGCACTCGGCAATGATGGAACGAATATTAG 1180
QY 1034 GACCCATACCAACACATGATTCAGAAACCAAGAAACGCAAGTATTTTACCATAACC 1093
Db 1181 GACCCATACCAACACATGATTCAGAAACCAAGAAACGCAAGTATTTTACCATAACC 1240
QY 1094 AGCTAGATTGGATGAACACAGTCTCTGCTGTAGATATGTTAGAGAGCGCTGCAACCGT 1153
Db 1241 AGCTAGATTGGATGAACACAGTCTCTGCTGTAGATATGTTAGAGAGCGCTGCAACCGT 1300
QY 1154 TGAAGGAATTTATGTTTGTATGATGAAGAACATGAGAAACTGTTTGCCTGTTTGGAA 1213
Db 1301 TGAAGGAATTTATGTTTGTATGATGAAGAACATGAGAAACTGTTTGCCTGTTTGGAA 1360
QY 1214 GAATGTTAGATATGATCAACTCAAGAAATTAATCTTGATGAAGCAATGAGCAATCCTT 1273
Db 1361 GAATGTTAGATATGATCAACTCAAGAAATTAATCTTGATGAAGCAATGAGCAATCCTT 1420
QY 1274 TCTTTGACTTATTAAGAAAGAAATGAAATGGAATCAGTGTCTTACTATATCTTCT 1333
Db 1421 TCTTTGACTTATTAAGAAAGAAATGAAATGGAATCAGTGTCTTACTATATCTTCT 1480
QY 1334 AGAAGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATT 1393
Db 1481 AGAAGAGATTACTTAAGACTGTGTCAGTCAACTAAACATTTCTAATATTTTGTAAACATT 1540
QY 1394 AAATTTTGTACAGTTAAGTAAATATTGTATGTTTGTATGTTTGTATCAATPAGCAATTAAC 1453
Db 1541 AAATTTTGTACAGTTAAGTAAATATTGTATGTTTGTATGTTTGTATCAATPAGCAATTAAC 1600
QY 1454 TTGTTAAGCAAGTGTGTTGATATGCTATAGAAATTAATTAATTTTCTTTT 1513
Db 1601 TTGTTAAGCAAGTGTGTTGATATGCTATAGAAATTAATTAATTTTCTTTT 1660

QY 1514 GAAATTACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATAAATGTGATTG 1573
Db 1661 GAAATTACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATAAATGTGATTG 1720
QY 1574 ATCTTGCCCTTTTGTACATGGAGGTCACCTCTGAAGTGATTTTGTGAGTAAAGGAAT 1633
Db 1721 ATCTTGCCCTTTTGTACATGGAGGTCACCTCTGAAGTGATTTTGTGAGTAAAGGAAT 1780
QY 1634 CTTGACTACTTTT 1645
Db 1781 CTTGACTACTTT 1792

RESULT 8
AAC81772
ID AAC81772 standard; DNA; 1323 BP.
XX
AC AAC81772;
XX
DT 23-FEB-2001 (first entry)
XX
DE Human beacon partial coding sequence SEQ ID NO: 7.
XX
KW Israeli sand rat; beacon; obesity; NIDDM; energy balance; diabetes;
KW ligand; human; db.
XX
OS Homo sapiens.
XX
PN WO200064931-A1.
XX
PD 02-NOV-2000.
XX
PF 19-APR-2000; 2000WO-AU000342.
XX
PR 23-APR-1999; 99AU-00009919.
PR 24-MAR-2000; 2000AU-00006454.
XX
PA (AUTO-) AUTOGEN PTY LTD.
XX
PI Collier G, Walder K, Zimmet P;
XX
WPI; 2000-687311/67.
XX
Ligand of beacon protein useful for treating obesity, anorexia, energy
imbalance, diabetes, metabolic syndrome, dyslipidemia, hypertension and
insulin resistance.
XX
Claim 4; Fig 2; 67pp; English.
XX
The present invention is related to the isolation of a ligand known as
beacon from the Israeli sand rat. Beacon is associated with the
regulation of energy balance, and the protein, its coding sequence and
analogues can be used in the treatment of diabetes, obesity, anorexia,
energy imbalance, metabolic syndrome, dyslipidaemia, hypertension and
insulin resistance. In addition, they can be used in agriculture to
produce leaner animals
XX
Sequence 1323 BP; 452 A; 199 C; 232 G; 440 T; 0 U; 0 Other;
XX
Query Match 57.7%; Score 1300.6; DB 3; Length 1323;
Best Local Similarity 98.9%; Pred. No. 4.9e-271;
Matches 1309; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 626 TTCAATTTGACCAATCAGGAGATGGGTATCAGATCTGCCAGTCAATAAATTTTAC 685
Db 1 TTCAATTTGACCAATCAGGAGATGGGTATCAGATCTGCCAGTCAATAAATTTTAC 60
QY 686 ATCAATAATAATTAACCATACAGATCTGAAGCTGAAATATTTTGTGAGAGTCTG 745
Db 61 ATCAATAATAATTAACCATACAGATCTGAAGCTGAAATATTTTGTGAGAGTCTG 120
QY 746 ACTATGATAGTCAATAATAATTTCTAAATGAAACGTGATGAACGCACACTGAAAAACAG 805
Db 746 ACTATGATAGTCAATAATAATTTCTAAATGAAACGTGATGAACGCACACTGAAAAACAG 805
```


Best Local Similarity 100.0%; Pred. No. 7.2e-268; Matches 1286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	14	GTCAATTATTAGAACGAAGTCCCTTGAATGAGCGAGATTATCGGGACCGGAGATACGTTG	73
Db	161	GTCAATTATTAGAACGAAGTCCCTTGAATGAGCGAGATTATCGGGACCGGAGATACGTTG	220
Qy	74	ACGAATACAGGAATGACTATCTGTGAAGGATATGTTCTAGACATTATCAGAGACATTG	133
Db	221	ACGAATACAGGAATGACTACTGTGAAGGATATGTTCTAGACATTATCAGAGACATTG	280
Qy	134	AAAGCGGTATCCGAATCCACTGCAAGTAAATCTTCAGTCCGACGAGAGACGAGTCCCTA	193
Db	281	AAAGCGGTATCCGAATCCACTGCAAGTAAATCTTCAGTCCGACGAGAGACGAGTCCCTA	340
Qy	194	AAAGGAAGCGCAATAGACACCTGTTCAAAGTCATCAGTCACTGTTCCAAAGACCCACGAAGGA	253
Db	341	AAAGGAAGCGCAATAGACACCTGTTCAAAGTCATCAGTCACTGTTCCAAAGACCCACGAAGGA	400
Qy	254	AAAGATCCAGGATATAGAGGATGATGAGGAGGTCACTTGATCTGTCAAAGTGGAGACG	313
Db	401	AAAGATCCAGGATATAGAGGATGATGAGGAGGTCACTTGATCTGTCAAAGTGGAGACG	460
Qy	314	TTCTAAGAGCAAGATATGAATCTGTGACACTTTTGGGTGAAGGACCTTTGGCAAGTTG	373
Db	461	TTCTAAGAGCAAGATATGAATCTGTGACACTTTTGGGTGAAGGACCTTTGGCAAGTTG	520
Qy	374	TAGAGTGCATTGATCATGGCATGATGGCATGTCATGTAGCAGTCAAAATCGTAAAAAATG	433
Db	521	TAGAGTGCATTGATCATGGCATGATGGCATGTCATGTAGCAGTCAAAATCGTAAAAAATG	580
Qy	434	TAGCCCGTTACCGTGAAGCAGCTCGTTCAGAAAACTCAAGTATTAGAGCAGCTTAATAGTA	493
Db	581	TAGCCCGTTACCGTGAAGCAGCTCGTTCAGAAAACTCAAGTATTAGAGCAGCTTAATAGTA	640
Qy	494	CTGATCCCAATAGTGTCTTCGATGTGTCCAGATGTCTAGATGTGTAGATGTTGATCATCGTGC	553
Db	641	CTGATCCCAATAGTGTCTTCGATGTGTCCAGATGTCTAGATGTGTAGATGTTGATCATCGTGC	700
Qy	554	ATGTTTCTGATTGTCTTTGAACTACTCTGGGACTTTAGTACTTACGATTTCAATTAAGAAAAACA	613
Db	701	ATGTTTCTGATTGTCTTTGAACTACTCTGGGACTTTAGTACTTACGATTTCAATTAAGAAAAACA	760
Qy	614	GCTTTCTGCCATTTCAAATTGACCACATCAGGACAGATGGCGTATCAGATCTGCCAGTCAA	673
Db	761	GCTTTCTGCCATTTCAAATTGACCACATCAGGACAGATGGCGTATCAGATCTGCCAGTCAA	820
Qy	674	TAAATTTTTTACATCATATAATAATTAACCCATAAGATCTGAAGCCTTGAAAAATTTTGT	733
Db	821	TAAATTTTTTACATCATATAATAATTAACCCATAAGATCTGAAGCCTTGAAAAATTTTGT	880
Qy	734	TTGTGAGTCTGACTATGATGTAATTAATTTCTAAATGAACGCTGATGAAGCAGCAC	793
Db	881	TTGTGAGTCTGACTATGATGTAATTAATTTCTAAATGAACGCTGATGAAGCAGCAC	940
Qy	794	TGAAAAACACAGATATCAAAGTGTGTGACTTTTGGAAAGTGCAACGTATGATGAACATC	853
Db	941	TGAAAAACACAGATATCAAAGTGTGTGACTTTTGGAAAGTGCAACGTATGATGAACATC	1000
Qy	854	ACAGTATTTTGGTGTCTACCGGCACTACAGAGTCCCGAGGTCAATTTTGGCTTTAGGTT	913
Db	1001	ACAGTATTTTGGTGTCTACCGGCACTACAGAGTCCCGAGGTCAATTTTGGCTTTAGGTT	1060
Qy	914	GGTCTCAGCCTTTGATGTTTGGAGCATAGTTTGCATCTTATTGAATATTACCTTGGTT	973
Db	1061	GGTCTCAGCCTTTGATGTTTGGAGCATAGTTTGCATCTTATTGAATATTACCTTGGTT	1120
Qy	974	TCAGTCTTTTCAGACTCATGATAGTAAAGAGCACCTTGGCAATCATCGAAATATTAG	1033
Db	1121	TCAGTCTTTTCAGACTCATGATAGTAAAGAGCACCTTGGCAATCATCGAAATATTAG	1180
Qy	1034	GACCCATACCAACACATGATTTCAGAAAAACAGAAAAACGCAAGTATTTTACCATAACC	1093

expressing the protein is useful for producing a pharmaceutical composition. Compositions, antibodies and antisense molecules are useful for the treating a disease associated with STAT6 activation such as allergic diseases, inflammation, autoimmune diseases, diabetes, hyperlipidaemia, infections disease and cancers. Compositions are useful for treating disease associated with STAT6 activation and/or prevention of Th1 hyperactive diseases. Compositions are also useful in rheumatoid arthritis, osteoarthritis, systemic lupus erythematosus, sepsis, asthma, allergic rhinitis, ischaemic heart diseases, subarachnoid haemorrhage, viral hepatitis and AIDS. The protein has efficient promoting STAT6 activity. The protein or nucleic acid is effectively useful for screening compounds for treating and preventing disease associated with excessive activation or inhibition of STAT6. The present sequence represents a human cDNA encoding a protein which promotes STAT6 activation.

Sequence 1881 BP; 570 A; 344 C; 450 G; 517 T; 0 U; 0 Other; 0

Query Match 56.6%; Score 1277.4; DB 12; Length 1881;
Best Local Similarity 87.9%; Pred. No. 5.6e-266;
Matches 1413; Conservative 0; Mismatches 192; Indels 3;

Qy	14	GTCA	TATTTAGAA	GCAAGTCTCTTGAATGAGCGAGATTATCGGACCGGAGATACGTTG	73
Db	250				309
		GTCACTATTTAGAA	GCAAGATGCTTGAATGAGAGAGATTATCGGACCGGAGATACATTG		
Qy	74	ACGAATACAGGA	TGACTACTGTGCAAGGATATGTTCTTAGACATATTATCAGAGAGACATTG	133	
Db	310				369
		ATGAATACAGAA	TGACTCTGCAAGGATATGTTCCAGACATTACCATAGAGACGTTG		
Qy	134	AAAGCGGGTATCGAA	TCCACTGCAAGTAAATCTTTCAGTCCGACGAGGAGAAGCAGTCTTA	193	
Db	370				429
		AAAGCACTTACCGGATCCATTCGAGTAAATCCTCAGTCAGGACGAGAGAAGCAGCCCTA			
Qy	194	AAAGGAAGCGCA	TATAGACACTGTTCAGTCACTCATGTCACGTTCGAGAGAGCCACGGAAGGA	253	
Db	430				489
		AGAAAGCGTAATAGACCCCTGTGCAAGTCATCAGTCGCATTTGGAAGAGCACCGAAGGA			
Qy	254	AAAGATCCAGGAGTATAGAG	TGATGAGAGGGGTACCTGATCTGTCAAAGTGGAGACG	313	
Db	490				549
		AAAGATCCAGGAGTATAGAGATGATGAGAGGGGTACCTGATCTGTCAAAGTGGAGACG			
Qy	314	TTCTTAGAGCAAGATATGA	AATCGTGACACTTTTGGGTGAAGGAGCCTTTGGCAAAAGTTG	373	
Db	550				609
		TTCTTAGAGCAAGATATGA	AATCGTGACACTTTTAGGTGAAGGAGCCTTTGGCAAAAGTTG		
Qy	374	TAGAGTCATTTGATCGCA	TGGATGGCATGATGTAGCAGTGAATAATCGTAAAAAATG	433	
Db	610				669
		TAGAGTCATTTGATCACGGATGGATGGCTTTACATGTAGCAGTGAATAATTTGAAAAAATG			
Qy	434	TAGCCGTTACCGTGAAG	CAGCTCGTTTCAGAAATCCCAAGTATTAGACACCTTTAAATAGTA	493	
Db	670				729
		TAGCCGTTTACCGGAGGAGCTCGTTCTGAAATCCCAAGTATTGGAGCATTTGAACAGCA			
Qy	494	CTGATCCCAATAGTGTCT	TCGATGTGTCCAGATGCTAGAAATGGTTTGATCATCATGGTC	553	
Db	730				789
		CTGACCCCAACAGTGTCTTCCGATGCGTCAGATGCTAGAGTGGTTTGATCATCATGGTC			
Qy	554	ATGTTTGTAATGTGTTTGA	ACTACTGGGACTTAGTACTTACGATTTCAATTAAAGAAAAACA	613	
Db	790				849
		ATGTTTGTAATGTGTTTGA	CTGCTGGGACTTAGTACTATGATTTTATTAAAGAAAAATA		
Qy	614	GCITTTCTGCCAATTCAA	ATTGACCACATCAGGAGATGGCGTATCAGATCTGCCAGTCAA	673	
Db	850				909
		GTITTTCTGCCAATTCAA	ATTGATCACATCAGGCAAAATGGCTTATCAGATCTGCCAGTCTA		
Qy	674	TAAATTTTTTACATCAT	AATAAATTAACCCCATACAGATCTGAAGCCCTGAAAAATATTTTGT	733	
Db	910				969
		TAAATTTTTTACATCAT	AATAAATTAACACACACGAGACTTAAACCTGAAAAATATTTTAT		
Qy	734	TTGTGAAGCTGACTATGT	AGTCAAAATATAATTTCTAAAAATGAACGCTGATGAACGCAAC	793	
Db	970				1029
		TTGTGAAGCTGACTATGT	AGTCAAAATACAATTTCTAAATGAACACGAGATGAGCGCAT		

OS Homo sapiens.
FN WO200056753-A1.
XX
XX
PD 28-SEP-2000.
XX
XX
PF 16-MAR-2000; 2000WO-US006765.
XX
XX 23-MAR-1999; 99US-0126051P.
PR 10-DEC-1999; 99US-0169906P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Komatsoulis G;
PI WPI; 2000-594570/56.
XX
XX P-PSDB; AAB33724.
DR
XX
XX Nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating a disorder.
XX
XX Claim 1; Page 349-350; 410pp; English.
XX
XX The invention relates to the isolation of genes AAC59277-C59325 encoding
CC 49 human secreted proteins AAB33718-B33764. The genes can be used to
CC generate fusion proteins by linking to the gene for the human
CC immunoglobulin G Fc portion (SEQID1) for increasing the stability of the
CC fusion protein as compared to the human protein only. The genes and
CC proteins are useful for preventing, ameliorating or treating medical
CC conditions, e.g. by protein or gene therapy. The genes are isolated from
CC a range of human tissues disclosed in the specification. The nucleic
CC acids, proteins, antibodies and (ant)agonists are useful in the
CC diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer, and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b)
CC immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections
XX
SQ Sequence 1222 BP; 407 A; 194 C; 229 G; 392 T; 0 U; 0 Other;

Query Match 52.5%; Score 1185; DB 3; Length 1222;
Best Local Similarity 99.9%; Pred. No. 4,6e-246;
Matches 1196; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 446 GTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTACTGATCCCAATA 505
Db 9 GTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTACTGATCCCAATA 68

QY 506 GTGCTTCGGATGTCAGATGCTAGAGTGGTTGATCATCATGCTGATGTTGTATTG 565
Db 69 GTGCTTCGGATGTCAGATGCTAGAGTGGTTGATCATCATGCTGATGTTGTATTG 128

QY 566 TGTTTGAAGTCTGGGACTTAGTACTAGCTTTTCATTAAGAAACAGCTTCTGCGCAT 625
Db 129 TGTTTGAAGTCTGGGACTTAGTACTAGCTTTTCATTAAGAAACAGCTTCTGCGCAT 188

QY 626 TTCAAATTGACCACATCAGGATGGGATATCAGATCTGCCAGTCAATAAATTTTAC 685
Db 189 TTCAAATTGACCACATCAGGATGGGATATCAGATCTGCCAGTCAATAAATTTTAC 248

QY 686 ATCATAATAAATTAACCATCAGATCTGAAGCCTGAAATAATTTTGTGTAAGTCTG 745
Db 249 ATCATAATAAATTAACCATCAGATCTGAAGCCTGAAATAATTTTGTGTAAGTCTG 308

QY 746 ACTATGATGCAAAATATATCTTAAATGAACGATGAGCGACACCTGAAACACAG 805
Db 309 ACTATGATGCAAAATATATCTTAAATGAACGATGAGCGACACCTGAAACACAG 368

QY 806 ATATCAAAAGTTGTTGACTTTGGAAGTGCAACGATATGATGATGAACATCACAGTACTTTGG 865
Db 369 ATATCAAAAGTTGTTGACTTTGGAAGTGCAACGATATGATGATGAACATCACAGTACTTTGG 428

QY 866 TGTCTACCCGGCACTACAGAGCTCCGAGAGTCATTTTGGCTTTAGTTGGTCTCAGCCTT 925
Db 429 TGTCTACCCGGCACTACAGAGCTCCGAGAGTCATTTTGGCTTTAGTTGGTCTCAGCCTT 488

QY 926 GTGATGTTTGGAGCATAGTTGCAATCTTATTTGAATATTTACCTGGTTTCAAGTCTTTC 985
Db 489 GTGATGTTTGGAGCATAGTTGCAATCTTATTTGAATATTTACCTGGTTTCAAGTCTTTC 548

QY 986 AGACTCATGATAGTAAGAGCAGCTGCAATGATGGAACGAATATTAGACCCATACAC 1045
Db 549 AGACTCATGATAGTAAGAGCAGCTGCAATGATGGAACGAATATTAGACCCATACAC 608

QY 1046 AACACATGATTTCAGAAAACGAAGAAACGCAAGTATTTTTCACATACACAGCTAGATTGGG 1105
Db 609 AACACATGATTTCAGAAAACGAAGAAACGCAAGTATTTTTCACATACACAGCTAGATTGGG 668

QY 1106 ATGAACACAGTTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGTTGAAGGAATTTA 1165
Db 669 ATGAACACAGTTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGTTGAAGGAATTTA 728

QY 1166 TGTCTTGTGATGATGAAGAAACATGAGAAACCTGTTTACCTGGTTGCAAGAAATTTAGAA 1225
Db 729 TGTCTTGTGATGATGAAGAAACATGAGAAACCTGTTTACCTGGTTGCAAGAAATTTAGAA 788

QY 1226 ATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTCAGCATCTCTTTCTTTGACTTAT 1285
Db 789 ATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTCAGCATCTCTTTCTTTGACTTAT 848

QY 1286 TAAAAAGAAATGAATGGGAATCAGTGGTCTTACTATATATATCTTCTTAGAAGAGATTAC 1345
Db 849 TAAAAAGAAATGAATGGGAATCAGTGGTCTTACTATATATATCTTCTTAGAAGAGATTAC 908

QY 1346 TTAAGACTGTGTCAGTCACTAAACATCTTAATATTTTGTAAACATTAATTTTGT 1405
Db 909 TTAAGACTGTGTCAGTCACTAAACATCTTAATATTTTGTAAACATTAATTTTGT 968

QY 1406 ACAGTAAAGTCAATATTTGATGTTTGTATCAATAGCAATTAATTAATCTGTTAAGCAAG 1465
Db 969 ACAGTAAAGTCAATATTTGATGTTTGTATCAATAGCAATTAATTAATCTGTTAAGCAAG 1028

QY 1466 TATGGTCTTGTATATGCAATTAAGAAATTAATAATTTTCTTTTGTAAATTAACCAT 1525
Db 1029 TATGGTCTTGTATATGCAATTAAGAAATTAATAATTTTCTTTTGTAAATTAACCAT 1088

QY 1526 TTTAAATACCTTTGAAATAT - CCTTTGTGTCAGTGAATAATGATGATGATCTTGCCTTT 1584
Db 1089 TTTAAATACCTTTGAAATATATACCTTTGTGTCAGTGAATAATGATGATGATCTTGCCTTT 1148

QY 1585 TGTACATGGAGGTCACTCTCAAGTATTTTGTAGTAAAGAAATCTTGACTA 1641
Db 1149 TGTACATGGAGGTCACTCTCAAGTATTTTGTAGTAAAGAAATCTTGACTA 1205

RESULT 12
ADI31373
ID ADI31373 standard; cDNA; 1456 BP.
XX
XX ADI31373;
XX AC
XX XX
XX DT 17-JUN-2004 (first entry)
XX
XX DE Human cDNA #699.
XX
XX KW Human; gene; ss; immunological response; immunopathological condition;
KW Crohn's disease; asthma; ulcerative colitis; hypereosinophilia;
KW irritable bowel syndrome; osteoarthritis; rheumatoid arthritis;
KW acute monocytic leukaemia; antiinflammatory; antisthmatic; antiulcer;
KW osteopathic; antiarthritic; antirheumatic; cyostatic.
XX

OS Homo sapiens.
PN US6607879-B1.
XX
PD 19-AUG-2003.
XX
PF 09-FEB-1998; 98US-00023655.
XX
PR 09-FEB-1998; 98US-00023655.
XX
PA (INCY-) INCYTE CORP.
XX
PI Cocks BG, Stuart SG, Seilhamer JJ;
XX
DR WPI; 2003-895307/82.
XX
PT A composition comprising a plurality of cDNAs, useful for detecting
PT altered expression of genes in an immunological response or for
PT diagnosing and treating an immunopathology, e.g. Crohn's disease, asthma
PT or osteoarthritis.
XX
PS Claim 1; SEQ ID NO 699; 50pp; English.
XX
CC The invention relates to a composition comprising a plurality of cDNAs
CC for detecting the altered expression of genes in an immunological
CC response. The invention also relates to a method of diagnosing or
CC monitoring the treatment of an immunopathological condition in a sample,
CC comprising obtaining nucleic acids from a sample, contacting the nucleic
CC acids of the sample with an array comprising the plurality of cDNAs under
CC conditions to form one or more hybridisation complexes, detecting the
CC hybridisation complexes and comparing the levels of the detected
CC hybridisation complexes with the level of hybridisation complexes
CC detected in a non-diseased sample, where an altered level of the detected
CC hybridisation complexes correlates with the presence of an
CC immunopathological condition. Also disclosed are an expression profile
CC comprising a microarray and a plurality of detectable complexes and a
CC method for identifying a plurality of polynucleotide probes. The cDNAs
CC are useful as hybridisable array elements in a microarray for monitoring
CC the expression of target polynucleotides. The microarray can be used in
CC the diagnosis of an immunopathology, such as Crohn's disease, asthma,
CC ulcerative colitis, hypereosinophilia, irritable bowel syndrome,
CC osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in
CC identifying agents for the treatment of the diseases. The microarray may
CC also be used in drug discovery and development, toxicological and
CC carcinogenicity studies, forensics or pharmacogenomics. The composition
CC may also be used in purification of a subpopulation of mRNAs, cDNAs or
CC genomic fragments. This sequence represents a human cDNA of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification but was obtained in electronic format directly
CC from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 1456 BP; 480 A; 213 C; 261 G; 502 T; 0 U; 0 Other;

Query Match 52.2%; Score 1178.2; DB 11; Length 1456;
Best Local Similarity 99.7%; Pred. No. 1.4e-244;
Matches 1191; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 932 TTTCGAGCATAGTTGCATCTTATTGAATATTACCTGGTTTTCAGTCTTTCAGACTC 991
DB 262 TGTAGCGCATAGTTGCATCTTATTGAATATTACCTGGTTTTCAGTCTTTCAGACTC 321
QY 992 ATGATAGTAAGAGACCTGGCAATGATGGACGAATATTAGGACCCATACCACACACA 1051
DB 322 ATGATAGTAAGAGACCTGGCAATGATGGACGAATATTAGGACCCATACCACACACA 381
QY 1052 TGATTCAGAAAACAGAAAACCGCAGTATTTTCACCAATACCAGTAGATGGGATGAAC 1111
DB 382 TGATTCAGAAAACAGAAAACCGCAGTATTTTCACCAATACCAGTAGATGGGATGAAC 441
QY 1112 ACAGTTCTGCTGGTAGATATGTTAGGACGCTGCAAAACCGTTGAAGGAATTTATGCTTTT 1171
DB 442 ACAGTTCTGCTGGTAGATATGTTAGGACGCTGCAAAACCGTTGAAGGAATTTATGCTTTT 501

QY 1172 GTCATGATGAAGAACATGAGAAACCTGTTGACCTGGTTCGAGAAATGTTAGAAATATGATC 1231
DB 502 GTCATGATGAAGAACATGAGAAACCTGTTGACCTGGTTCGAGAAATGTTAGAAATATGATC 561
QY 1232 CAACCTCAAAGAAATTACCTTGGATGAAGCAATTCGAGCATCCTTTCTTTGACATTTATAAAA 1291
DB 562 CAACCTCAAAGAAATTACCTTGGATGAAGCAATTCGAGCATCCTTTCTTTGACATTTATAAAA 621
QY 1292 AGAAATGAAATGGGAATCAGTGGTCTTACTATATATCTCTCTAGAAGAGATTACTTAAGA 1351
DB 622 AGAAATGAAATGGGAATCAGTGGTCTTACTATATATCTCTCTAGAAGAGATTACTTAAGA 681
QY 1352 CTGTGTCAGTCAACTAAACATTCCTAAATATTTTGTGTAACATTAATATTTTGTACAGTT 1411
DB 682 CTGTGTCAGTCAACTAAACATTCCTAAATATTTTGTGTAACATTAATATTTTGTACAGTT 741
QY 1412 AAGTGTAAATATCTATGTTTGTATCAATAGCATAATTAACCTTGTGTAAGCAAGTATGCT 1471
DB 742 AAGTGTAAATATCTATGTTTGTATCAATAGCATAATTAACCTTGTGTAAGCAAGTATGCT 801
QY 1472 CTTGATATATGCAATAGAAAATTTAAATTTTCTTTTGGAAATTA-CCATTTTAA 1530
DB 802 CTTGATATATGCAATAGAAAATTTAAATTTTCTTTTGGAAATTTATCCATTTTAA 861
QY 1531 ATACCTTTGAAATATCCTTTGTGTCAGTGAATAATGTGATGATCTTGCCTTTTGTAACA 1590
DB 862 ATACCTTTGAAATATCCTTTGTGTCAGTGAATAATGTGATGATCTTGCCTTTTGTAACA 921
QY 1591 TGGAGTCACTCTGAAAGTGAATTTTGTGAGTAAAGGAAATCTTGACTACTTTATATAT 1650
DB 922 TGGAGTCACTCTGAAAGTGAATTTTGTGAGTAAAGGAAATCTTGACTACTTTATATAT 981
QY 1651 CTTAAAGGAATATCTTTTATATATCTTCAAAATTTAGAACTTAACTTTAAAGTTTCTTC 1710
DB 982 CTTAAAGGAATATCTTTTATATATCTTCAAAATTTAGAACTTAACTTTAAAGTTTCTTC 1041
QY 1711 TGTAAATTTGAAAGGATGATTTATTAATCTAGATAACAGGACTAGAAACCAAAA 1770
DB 1042 TGTAAATTTGAAAGGATGATTTATTAATCTAGATAACAGGACTAGAAACCAAAA 1101
QY 1771 CTCAGAAATGTTTACCTGTTAGAAATCTTATTAATTTTAACTGTTGTTATCTTTTTCATT 1830
DB 1102 CTCAGAAATGTTTACCTGTTAGAAATCTTATTAATTTTAACTGTTGTTATCTTTTTCATT 1161
QY 1831 GGGTGTGTCAGGGTGTAAACAGACATTCATGGAAGGCAATGCAGTTTGTCCATTTGTA 1890
DB 1162 GGGTGTGTCAGGGTGTAAACAGACATTCATGGAAGGCAATGCAGTTTGTCCATTTGTA 1221
QY 1891 CAGTTTGTGTTAAATAAACACCATACACACTTTTATTAAGATTAAATCTAACTGGAAGT 1950
DB 1222 CAGTTTGTGTTAAATAAACACCATACACACTTTTATTAAGATTAAATCTAACTGGAAGT 1281
QY 1951 CAGCTTGGAAAATGAGACATTTCCAAAGTATGTTTGGTGAATCAGATATAAATAATAGAAA 2010
DB 1282 CAGCTTGGAAAATGAGACATTTCCAAAGTATGTTTGGTGAATCAGATATAAATAATAGAAA 1341
QY 2011 TTCTGATGAGAGGTTTCAGTTTAAATACCAAGTCTTTAGGAGTCTTAACTGTCAGC 2070
DB 1342 TTCTGATGAGAGGTTTCAGTTTAAATACCAAGTCTTTAGGAGTCTTAACTGTCAGC 1401
QY 2071 ATCTGTTTATCAATGACATAAATACGTAACCTTATTAAGATTAAATTAAGTTTATTAAT 2125
DB 1402 ATCTGTTTATCAATGACATAAATACGTAACCTTATTAAGATTAAATTAAGTTTATTAAT 1456

RESULT 13
ADI26144
ID ADI26144 standard; cDNA; 1865 BP.
XX
AC ADI26144;
XX
DT 22-APR-2004 (first entry)
XX


```
Db 1181 GACCCATCCAGCACATATGATCCAGACAGAGAAACGCAAGTATTTCACCATACCC 1240
QY 1094 ACCTAGATTGGGATGAACACAGTTCTGCTGGTAGATATGTTAGGAGACGCTGCAAAACCGT 1153
Db 1241 ACCTAGATTGGGACGAGCATAGTTGAGCTGGGAGATATGTTAGGAGACGCTGCAAGCCGT 1300
QY 1154 TGAAGGATTTATGCTTTGTCATGATGAGAACATGAGAACTGTTGACCTGGTTCGAA 1213
Db 1301 TAAAGGAATTTATGCTGTCATGACGAAGACATGAGAACTGTTGACCTGGTTCGAA 1360
QY 1214 GAATGTTAGATATATCCAACTCAAAGAAATTACCTTGGATCAAGCATTCGAGCATCCCT 1273
Db 1361 GAATGTTGAGATATGACCCAGGAGGAGGATCACCTTGGATCAAGCATTCGAGCATCCCT 1420
QY 1274 TCTTTGACTTATAAAAAAGAAATGA 1299
Db 1421 TCTTTGACTTATAAAAAAGAAATGA 1446

RESULT 15
AAD32039
ID AAD32039 standard; DNA; 21234 BP.
AC AAD32039;
XX
DT 18-JUN-2002 (first entry)
DE Human kinase protein gene.
KW Human; kinase protein; enzyme; cytostatic; osteopathic; gene expression;
KW colon-moderately differentiated adenocarcinoma; chromosome mapping; gene;
KW drug screening; therapeutic; gene therapy; tissue typing; chromosome 5;
KW bone osteosarcoma; single nucleotide polymorphism; SNP; ds.
XX
OS Homo sapiens.
FH
FH Key Location/Qualifiers
FT variation replace(75..76,A)
FT exon 2007..2059 /standard_name= "Single nucleotide polymorphism"
FT intron /tag= b
FT 2060..3118 /tag= c
FT exon 3119..3341 /tag= d
FT intron 3342..4462 /tag= e
FT exon 4463..4553 /tag= f
FT intron 4554..4948 /tag= g
FT exon 4949..5015 /tag= h
FT intron 5016..8054 /tag= i
FT variation replace(7980,C) /tag= j
FT /standard_name= "Single nucleotide polymorphism"
FT exon 8055..8171 /tag= k
FT intron 8172..8258 /tag= l
FT exon 8259..8425 /tag= m
FT intron 8426..9007 /tag= n
FT variation replace(8571,T) /tag= o
FT /standard_name= "Single nucleotide polymorphism"
FT exon 9008..9102 /tag= p
FT intron 9103..9352 /tag= q
FT exon 9353..9482 /tag= r
FT intron 9483..13437 /tag= s
FT variation replace(11257,Y) /tag= t
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(11684,C) /tag= u
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(13312,C) /tag= v
FT /standard_name= "Single nucleotide polymorphism"
FT exon 13438..13520 /tag= w
FT intron 13521..16514 /tag= x
FT exon 16515..16594 /tag= y
FT intron 16595..17962 /tag= z
FT variation replace(17110,T) /tag= aa
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(17451,C) /tag= ab
FT exon 17963..18053 /tag= ac
FT intron 18054..18139 /tag= ad
FT exon 18140..18280 /tag= ae
FT variation replace(20766,G) /tag= af
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(20913..20914,T) /tag= ag
FT /standard_name= "Single nucleotide polymorphism"
XX WO200216567-A2.
XX 28-FEB-2002.
XX 24-AUG-2001; 2001WO-US026389.
XX 24-AUG-2000; 2000US-0227470P.
XX 19-MAR-2001; 2001US-00810671.
XX (APPL-) APPLERA CORP.
XX Yan C, Ye J, Ketchum KA, Di Francesco V, Beasley EM;
XX WPI; 2002-269354/31.
XX
XX New human kinase proteins and nucleic acids, useful in drug screening
XX assays, identifying modulators of kinase activity or treating disorders
XX characterized by absence or unwanted expression of the protein.
XX
XX Claim 4; Fig 3; 8lpp; English.
XX
XX The invention relates to isolated human kinase proteins and nucleic
XX acids. The nucleic acid and peptide sequences can be used as models for
XX the development of human therapeutic targets, aid in the identification
XX of therapeutic proteins and serve as targets for the development of human
XX therapeutic agents that modulate kinase activity in cells and tissues
XX that express the kinase. The nucleic acids are useful as probes or
XX primers, in constructing recombinant vectors, for expressing antigenic
XX portions of the proteins, chromosome mapping, drug screening, testing an
XX individual for a genotype, and for gene therapy in patients containing
XX cells that are aberrant in kinase gene expression. The proteins may be
XX used in drug screening assays, in the identification of compounds that
```

CC modulate, stimulate or inhibit kinase activity, in pharmacogenomic
CC analysis, in treating disorders characterised by an absence or unwanted
CC expression of the protein (bone osteosarcoma, or colon-moderately
CC differentiated adenocarcinoma), and in generating antibodies specific for
CC the peptides. Such antibodies can be used to detect the protein in situ,
CC in vitro, or in cell lysate or supernatant, to isolate and purify the
CC proteins from host cells, pharmacogenomic analysis, tissue typing, and in
CC inhibiting protein function. The present sequence is human kinase protein
CC gene located on chromosome 5
XX
SQ Sequence 21234 BP; 6065 A; 3772 C; 4258 G; 7139 T; 0 U; 0 Other;

Query Match 48.6%; Score 1097; DB 6; Length 21234;
Best Local Similarity 100.0%; Pred. No. 1e-226;
Matches 1097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1157 AGGAATTTATGCTTTGTCATGATGAGAAACATGAGAACTGTTGACCTGGTTCGAAGAA 1216
DB 18138 AGGAATTTATGCTTTGTCATGATGAGAAACATGAGAACTGTTGACCTGGTTCGAAGAA 18197

QY 1217 TGTTAGATATGATCCAACTCAAGAAATTTACCTGGATGAAGCATTTGAGCATCTCTTCT 1276
DB 18198 TGTTAGATATGATCCAACTCAAGAAATTTACCTGGATGAAGCATTTGAGCATCTCTTCT 18257

QY 1277 TTGACTTATTAAGAAATGAAATGGAATCAAGTGGTCTTACTATATCTCTCTAGA 1336
DB 18258 TTGACTTATTAAGAAATGAAATGGAATCAAGTGGTCTTACTATATCTCTCTAGA 18317

QY 1337 AGAGATTACTTAAGCTGTGTCAGTCAACTAAGCATTTCTAATATTTTGTAAACATTAAA 1396
DB 18318 AGAGATTACTTAAGCTGTGTCAGTCAACTAAGCATTTCTAATATTTTGTAAACATTAAA 18377

QY 1397 TTATTTGTGACAGTTAAGTGTAAATATTGTATGTTTGTATCAATAGCATTAATTAACCTTG 1456
DB 18378 TTATTTGTGACAGTTAAGTGTAAATATTGTATGTTTGTATCAATAGCATTAATTAACCTTG 18437

QY 1457 TTAAGCAAGTATGGTCTTGATAATGCAATTAGAAAATTTAAATTTATTTCTTTTGAA 1516
DB 18438 TTAAGCAAGTATGGTCTTGATAATGCAATTAGAAAATTTAAATTTATTTCTTTTGAA 18497

QY 1517 ATTACCATTTTAAATACCTTTGAAATATCTTTGTGTCAGTGATTAATGTGATGATC 1576
DB 18498 ATTACCATTTTAAATACCTTTGAAATATCTTTGTGTCAGTGATTAATGTGATGATC 18557

QY 1577 TTGCCTTTTGTACATGGAGGTCACTCTGAAGTGAATTTTTTTGAGTAAAGGAATCTT 1636
DB 18558 TTGCCTTTTGTACATGGAGGTCACTCTGAAGTGAATTTTTTTGAGTAAAGGAATCTT 18617

QY 1637 GACTACTTTATATCTTTAAAGGAATATTCTTTATATATCTTCAAAATTAGAACTTAACCTTT 1696
DB 18618 GACTACTTTATATCTTTAAAGGAATATTCTTTATATATCTTCAAAATTAGAACTTAACCTTT 18677

QY 1697 AAAAGTTTTCTCTGTAAATTTGTGAACGGGTGATTTATTAATTAATTAAGTGAAGCATG 1756
DB 18678 AAAAGTTTTCTCTGTAAATTTGTGAACGGGTGATTTATTAATTAATTAAGTGAAGCATG 18737

QY 1757 ACTAGAAACCAAACTCAGAAAATCTTTACTGTTAGAAATTTCTATTAAATTTTAAAGTGTG 1816
DB 18738 ACTAGAAACCAAACTCAGAAAATCTTTACTGTTAGAAATTTCTATTAAATTTTAAAGTGTG 18797

QY 1817 TATCTTTTTCATTCGGGTGATGTCAGGGTGATPAACCCAGCATTTCAATGAAAGGCATGCAG 1876
DB 18798 TATCTTTTTCATTCGGGTGATGTCAGGGGTGATPAACCCAGCATTTCAATGAAAGGCATGCAG 18857

QY 1877 TTTGTCCATTGTGACAGTTTGTAAATTAACCAACATACACACTTTTATTAAGATTAAAA 1936
DB 18858 TTTGTCCATTGTGACAGTTTGTAAATTAACCAACATACACACTTTTATTAAGATTAAAA 18917

QY 1937 TCTAACTGGAAAGTCAGCTTGGAAAATGACATTTCCAAAGTATGTTTGGTGAGTCAACAGA 1996
DB 18918 TCTAACTGGAAAGTCAGCTTGGAAAATGACATTTCCAAAGTATGTTTGGTGAGTCAACAGA 18977

QY 1997 TATAAAAATAGAAATTTCTGATGAGAGGTTTCAGTTTTTAAATACCAAGTCCCTTAGGAGTCT 2056

Search completed: March 16, 2005, 09:01:21
Job time : 1200.6 secs

DB 18978 TATAAAAATAGAAATTTCTGATGAGAGGTTTCAGTTTTTAATACCAAGTCCCTTAGAGTCT 19037
QY 2057 TAACATTTGGCCAGCATCTGTTTATCAAAATGACATATAATACGTAACCTTATAAGAAATTAAG 2116
DB 19038 TRACATTTGGCCAGCATCTGTTTATCAAAATGACATATAATACGTAACCTTATAAGAAATTAAG 19097
QY 2117 TTTTATTAAATTAGGCAATTTTATGCTCTGATATAATTTCTTACGGAGAAAGAGGATTTGATTG 2176
DB 19098 TTTTATTAAATTAGGCAATTTTATGCTCTGATATAATTTCTTACGGAGAAAGAGGATTTGATTG 19157

QY 2177 GAAAGCAGTTTGGGAAGAAAGTCTGCTGAAATTTCCAGAAATTTAAATTGATTTGGTTACAT 2236
DB 19158 GAAAGCAGTTTGGGAAGAAAGTCTGCTGAAATTTCCAGAAATTTAAATTGATTTGGTTACAT 19217

QY 2237 AAACCTTTTGTGACTTTCAG 2253
DB 19218 AAACCTTTTGTGACTTTCAG 19234

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 16, 2005, 04:53:38 ; Search time 9703.74 Seconds
(without alignments)

11265.245 Million cell updates/sec

Title: US-10-801-671-1_COPY_72_2327

Perfect score: 2256

Sequence: 1 gacactcatccagtcattt.....aaacttttgacttcagaaa 2256

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_ats.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2256	100.0	2354	6 AR232168	AR232168 Sequence
2	2256	100.0	2354	6 AR406007	AR406007 Sequence
3	2256	100.0	2354	6 AR534243	AR534243 Sequence
4	2256	100.0	2354	6 AX392827	AX392827 Sequence
5	2048	90.8	2488	6 AX056410	AX056410 Sequence
6	2046.6	90.7	2446	6 AR077239	AR077239 Sequence
7	2046.6	90.7	2446	6 BD107903	BD107903 Protein k
8	1632	72.3	1814	9 AF294429	AF294429 Homo sapi
9	1629.4	72.2	1792	6 AX961896	AX961896 Sequence
10	1318.4	58.4	3492	9 BC063116	BC063116 Homo sapi
11	1284	56.9	1446	12 AX335726	AX335726 Synthetic
12	1277.6	56.6	1914	10 BC012675	BC012675 Mus muscu
13	1277.4	56.6	1881	6 AX961894	AX961894 Sequence
14	1185	52.5	1222	6 BD260616	BD260616 49 human
15	1178.2	52.2	1456	6 AR380154	AR380154 Sequence
16	1118.8	49.6	1865	6 AX961898	AX961898 Sequence
17	1097.4	48.6	1549	10 AF033566	AF033566 Mus muscu
18	1097	48.6	21234	6 AR232169	AR232169 Sequence
19	1097	48.6	21234	6 AR406008	AR406008 Sequence

20	1097	48.6	21234	6 AR534244	AR534244 Sequence
21	1097	48.6	21234	6 AX392829	AX392829 Sequence
c 22	1097	48.6	159266	2 AC026097	AC026097 Homo sapi
c 23	1097	48.6	170624	9 AC022096	AC022096 Homo sapi
24	1009.6	44.8	1445	10 BC002220	BC002220 Mus muscu
25	995	44.1	1150	6 CQ718110	CQ718110 Sequence
26	913.4	40.5	1414	10 AY653308	AY653308 Mus muscu
27	910	40.3	1416	10 BC079006	BC079006 Rattus no
28	880.8	39.0	1745	5 BX931976	BX931976 Gallus ga
29	824.4	36.5	1864	9 AB056421	AB056421 Macaca fa
30	814.8	36.1	1757	6 CQ715616	CQ715616 Sequence
31	813.2	36.0	1773	9 BC031549	BC031549 Homo sapi
32	811.6	36.0	1834	9 HUMCLK1B	L29219 Homo sapien
33	768.8	34.1	1705	10 MMCLKMRNA	X57186 Mouse clk m
34	765.6	33.9	1740	10 MUSSTYKIN	M38381 Mouse serin
35	692	30.7	1946	5 CR848121	CR848121 Xenopus t
36	677.8	30.0	1544	5 BX936001	BX936001 Gallus ga
37	669.4	29.7	1017	5 BX929601	BX929601 Gallus ga
38	644.4	28.6	1743	9 HUMCLK1A	L29222 Homo sapien
39	643.2	28.5	1750	9 HUMKINCD	M59287 Human prote
40	640.6	28.4	2254	6 AR531674	AR531674 Sequence
c 41	639.4	28.3	1535	11 BV179987	BV179987 sqm10751
42	630.6	28.0	1643	10 MUSMCLK	L29221 Mus musculu
43	628.2	27.8	4862	10 BC062887	BC062887 Mus muscu
44	617	27.3	3056	10 BC081942	BC081942 Rattus no
45	595.2	26.4	1634	5 BX929386	BX929386 Gallus ga

ALIGNMENTS

RESULT 1
AR232168
LOCUS AR232168 2354 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 1 from patent US 6455291.
ACCESSION AR232168
VERSION AR232168.1 GI:27274056
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2354)
AUTHORS Yan, C., Ye, J., Ketchum, K.A., Di Francesco, V. and Beasley, E.M.
TITLE Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins, and uses thereof
JOURNAL Patent: US 6455291-A 1 24-SEP-2002;
FEATURES
source Location/Qualifiers
1..2354
/organism="unknown"
/mol_type="genomic DNA"

Query Match	100.0%	Score	2256;	DB	6;	Length	2354;
Best Local Similarity	100.0%;	Pred. No.	0;	Mismatches	0;	Indels	0;
Matches	2256;	Conservative	0;				
Qy	1	GACACTCATCCAGTCATTATTTAGAACCAAGGTCCTTGAATGACGAGATTATCGGCAC	60				
Db	72	GACACTCATCCAGTCATTATTTAGAACCAAGGTCCTTGAATGACGAGATTATCGGCAC	131				
Qy	61	CGGAGATACGTTGACGAATACAGGAATCACTACTGTGTGAAGGATATGTTCTTAGACATTAT	120				
Db	132	CGGAGATACGTTGACGAATACAGGAATCACTACTGTGTGAAGGATATGTTCTTAGACATTAT	191				
Qy	121	CACAGACATTGAAGCGGGTATCGAATCCACTGCAGTAAATCTTCAGTCGGACGAGG	180				
Db	192	CACAGACATTGAAGCGGGTATCGAATCCACTGCAGTAAATCTTCAGTCGGACGAGG	251				
Qy	181	AGAACGAGTCCTAAAGAACGCGCAATAGACACTGTTTCAAGTCATCATGTCAGTTCGGAAG	240				
Db	252	AGAACGAGTCCTAAAGAACGCGCAATAGACACTGTTTCAAGTCATCATGTCAGTTCGGAAG	311				
Qy	241	AGCCACCGAAGAAAGATCCAGGATATAGAGGATGATGAGAGGGGTACCTGATCTGT	300				

Db 312 AGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATAGGAGGTCACCTGATCTGT 371
 QY 301 CAAGTGGAGAGCTTCTAAGAGCAAGATATGAATCGTGGACACTTTGGGTGAAGAGCC 360
 Db 372 CAAGTGGAGAGCTTCTAAGAGCAAGATATGAATCGTGGACACTTTGGGTGAAGAGCC 431
 QY 361 TTTGGCAAGTTGTAGAGTGAATGATCATGCGATGGAATGCGATGAGTGAAGTGA 420
 Db 432 TTTGGCAAGTTGTAGAGTGAATGATCATGCGATGGAATGCGATGAGTGAAGTGA 491
 QY 421 ATCGTAAAGAAATGTAGGCGCTTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAG 480
 Db 492 ATCGTAAAGAAATGTAGGCGCTTACCGTGAAGCAGCTCGTTCAGAAATCCAAAGTATTAG 551
 QY 481 CACTTAAATAGTACTGATCCCAATAGTGTCTTCGATGTCAGAGTCTAGAAATGGTTT 540
 Db 552 CACTTAAATAGTACTGATCCCAATAGTGTCTTCGATGTCAGAGTCTAGAAATGGTTT 611
 QY 541 GATCATCATGTCATGTTGTATGTTGTGTTGTAATCTGAACTGAGTGGAGCTTACGATTTT 600
 Db 612 GATCATCATGTCATGTTGTATGTTGTGTTGTAATCTGAACTGAGTGGAGCTTACGATTTT 671
 QY 601 ATTAAGAAACAGCTTTCTGCCATTTCCAAATTCGAATTCAGCAGATGCGGTATCAG 660
 Db 672 ATTAAGAAACAGCTTTCTGCCATTTCCAAATTCGAATTCAGCAGATGCGGTATCAG 731
 QY 661 ATCTGCCAGTCAATAAAATTTTACATCAATAAATTAATTAATTAATTAATTAATTAAT 720
 Db 732 ATCTGCCAGTCAATAAAATTTTACATCAATAAATTAATTAATTAATTAATTAATTAAT 791
 QY 721 GAAATATTTTGTGTAAGTCTGACATGATGATGATGATGATGATGATGATGATGATGAT 780
 Db 792 GAAATATTTTGTGTAAGTCTGACATGATGATGATGATGATGATGATGATGATGATGAT 851
 QY 781 GATGAACGACACTGAAAGAAACACAGATATCAAAATTCGAATTCAGCAGATGCGGTAT 840
 Db 852 GATGAACGACACTGAAAGAAACACAGATATCAAAATTCGAATTCAGCAGATGCGGTAT 911
 QY 841 GATGATGATCACTACAGTACTTGTGTCCTACCGGCACTACAGAGCTCCCGAGTCAAT 900
 Db 912 GATGATGATCACTACAGTACTTGTGTCCTACCGGCACTACAGAGCTCCCGAGTCAAT 971
 QY 901 TTGGCTTTAGTGTGCTCAGCTTGTGATGTTTGGAGCATAGGTTGCAATTTATTGAA 960
 Db 972 TTGGCTTTAGTGTGCTCAGCTTGTGATGTTTGGAGCATAGGTTGCAATTTATTGAA 1031
 QY 961 TATTACCTTTGGTTTCAAGTCTTTGAGTCTCATGATGATGATGATGATGATGATGATGAT 1020
 Db 1032 TATTACCTTTGGTTTCAAGTCTTTGAGTCTCATGATGATGATGATGATGATGATGATGAT 1091
 QY 1021 GAAAGAAATATTAGGACCCATACCAACACATGATTCAGAAACAAAGAAACGCAAGTAT 1080
 Db 1092 GAAAGAAATATTAGGACCCATACCAACACATGATTCAGAAACAAAGAAACGCAAGTAT 1151
 QY 1081 TTTTCAACATAACAGTACTAGATGGGATGAACACAGTCTTGTGTTAGATGATGATGATGAT 1140
 Db 1152 TTTTCAACATAACAGTACTAGATGGGATGAACACAGTCTTGTGTTAGATGATGATGATGAT 1211
 QY 1141 CGCTGCAAAACCGTTGAAGAAATTTATGTTTGTGATGATGATGATGATGATGATGATGAT 1200
 Db 1212 CGCTGCAAAACCGTTGAAGAAATTTATGTTTGTGATGATGATGATGATGATGATGATGAT 1271
 QY 1201 GACCTGTTTGAAGAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
 Db 1272 GACCTGTTTGAAGAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1331
 QY 1261 TTGCAGCATCTCTTTGACCTTATTAAGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 1320
 Db 1332 TTGCAGCATCTCTTTGACCTTATTAAGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 1391
 QY 1321 TATATATCTCTAGAGGATTAATTAAGATGATGATGATGATGATGATGATGATGATGATGAT 1380

Db 1392 TATATATCTCTAGAGGATTAATTAAGATGATGATGATGATGATGATGATGATGATGAT 1451
 QY 1381 TTTTGTAAACATTAATTTATTTTGTACAGTTAAAGTAAATATTGATGATGATGATGATGAT 1440
 Db 1452 TTTTGTAAACATTAATTTATTTTGTACAGTTAAAGTAAATATTGATGATGATGATGATGAT 1511
 QY 1441 TAGCATAAATTAATTTTGTAAAGCAAGTATGCTTTGATAATGCAATTAAGAAATTTAA 1500
 Db 1512 TAGCATAAATTAATTTTGTAAAGCAAGTATGCTTTGATAATGCAATTAAGAAATTTAA 1571
 QY 1501 AATTTTCTTTTGAAGAAATTAATTTTAAATACCTTTGAATATCTTTGCTGCTGCTGCTG 1560
 Db 1572 AATTTTCTTTTGAAGAAATTAATTTTAAATACCTTTGAATATCTTTGCTGCTGCTGCTG 1631
 QY 1561 ATAAATGTGATGATGATGCTTTTGTACATGAGAGTCACTCTGAAAGTGAATTTTGTG 1620
 Db 1632 ATAAATGTGATGATGATGCTTTTGTACATGAGAGTCACTCTGAAAGTGAATTTTGTG 1691
 QY 1621 AGTAAAGGAAATCTTGACTACTTTATTTATTTCTTAAAGAAATATTCTTTATATATCTT 1680
 Db 1692 AGTAAAGGAAATCTTGACTACTTTATTTATTTCTTAAAGAAATATTCTTTATATATCTT 1751
 QY 1681 TTTAGAACTTAACTTTAAAGTCTTTCTTCTGTAATTTGTTGAAACGGGTGATTTATTAT 1740
 Db 1752 TTTAGAACTTAACTTTAAAGTCTTTCTTCTGTAATTTGTTGAAACGGGTGATTTATTAT 1811
 QY 1741 CTCTAGATAAGCAGTACTAGAAACCAAACTCAGAAATCTTTACTGTTAGAAATTTCTAT 1800
 Db 1812 CTCTAGATAAGCAGTACTAGAAACCAAACTCAGAAATCTTTACTGTTAGAAATTTCTAT 1871
 QY 1801 TAAATTTTAAAGTGTATTTCTTTTCAATGGGTGATGATGAGGTGATGATGATGATGATGAT 1860
 Db 1872 TAAATTTTAAAGTGTATTTCTTTTCAATGGGTGATGATGAGGTGATGATGATGATGATGAT 1931
 QY 1861 ATGAAAGGCAATGAGTGTGCTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1920
 Db 1932 ATGAAAGGCAATGAGTGTGCTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1991
 QY 1921 TTAATTTAAAGTAAATCTAACTGAAAGTCACTGAAAGTCACTGAAAGTCACTGAAAGTCA 1980
 Db 1992 TTAATTTAAAGTAAATCTAACTGAAAGTCACTGAAAGTCACTGAAAGTCACTGAAAGTCA 2051
 QY 1981 TTTGGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
 Db 2052 TTTGGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2111
 QY 2041 AAGTCTTTAGGAGTCTTAAACATTCGAGCATCTGTTTATCAAATGACATAAATACGTAA 2100
 Db 2112 AAGTCTTTAGGAGTCTTAAACATTCGAGCATCTGTTTATCAAATGACATAAATACGTAA 2171
 QY 2101 ACCTATAGAAATTAAGTGTATTTAATTTAGGCAATTTATGCTGATGATGATGATGATGAT 2160
 Db 2172 ACCTATAGAAATTAAGTGTATTTAATTTAGGCAATTTATGCTGATGATGATGATGATGAT 2231
 QY 2161 AAAGAGGATTTGATTTGAAACAGTCTTGGGAAAGAGTCTGCTGCTGCTGCTGCTGCTGCTG 2220
 Db 2232 AAAGAGGATTTGATTTGAAACAGTCTTGGGAAAGAGTCTGCTGCTGCTGCTGCTGCTGCTG 2291
 QY 2221 AATTGATTTGTTTACATAAATCTTTTGTGCTTCCAGAAA 2256
 Db 2292 AATTGATTTGTTTACATAAATCTTTTGTGCTTCCAGAAA 2327

RESULT 2
 AR406007
 LOCUS AR406007 2354 bp DNA linear PAT 18-DEC-2003
 DEFINITION Sequence 1 from patent US 6630337.
 ACCESSION AR406007
 VERSION AR406007.1 GI:40155113
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.

REFERENCE	1 (bases 1 to 2354)
AUTHORS	Yan,C., Ye,J., Ketchum,K.A., Di Francesco,V. and Beasley,E.M.
TITLE	Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins, and uses thereof
JOURNAL	Patent: US 6630337-A 1 07-OCT-2003;
FEATURES	Location/Qualifiers
source	1..2354
	/organism= "unknown"
ORIGIN	/mol_type= "genomic DNA"
	Query Match 100.0%; Score 2256; DB 6; Length 2354;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 GACACTCATCCAGTCATTATTAGAACCAAGGTCCTTGAATGAGCGAGATTATCGGGAC 60
Db	72 GACACTCATCCAGTCATTATTAGAACNAGGTCCTTGAATGAGCGAGATTATCGGAC 131
Qy	61 CGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTCTCTAGACATTAT 120
Db	132 CGGAGATACGTTGACGAATACAGGAATGACTACTGTGAAGGATATGTTCTCTAGACATTAT 191
Qy	121 CACAGAGCATTGAAAGCGGTTATCGAATCCACTCCAGTAAATCTTTCAGTCCGACGAGG 180
Db	192 CACAGAGCATTGAAAGCGGTTATCGAATCCACTCCAGTAAATCTTTCAGTCCGACGAGG 251
Qy	181 AGAAGCAGTCCTAAAGGAAGCGCAATAGACACTGTTCAAAGTCATCAGTCACGTTTCGAAG 240
Db	252 AGAAGCAGTCCTAAAGGAAGCGCAATAGACACTGTTCAAAGTCATCAGTCACGTTTCGAAG 311
Qy	241 AGCCA CGGAAGAAAAGATCCAGGATGATAGAGGATGATGAGGAGGGTCACCTGATCTGT 300
Db	312 AGCCACCGAAGAAAAGATCCAGGATGATAGAGGATGATGAGGAGGGTCACCTGATCTGT 371
Qy	301 CAAAGTGGAGAGGTTCTAAGAGCAAGATATGAAATCGTGGACACCTTGGGTGAAGGAGCC 360
Db	372 CAAAGTGGAGAGGTTCTAAGAGCAAGATATGAAATCGTGGACACCTTGGGTGAAGGAGCC 431
Qy	361 TTTGCGCAAAGTTGTAGAGTGCAATCATCATGCGATGGATGGCATCGATGTACGAGTAA 420
Db	432 TTTGCGCAAAGTTGTAGAGTGCAATCATCATGCGATGGATGGCATCGATGTACGAGTAA 491
Qy	421 ATCGTAAAAAATGTAGGCGGTTACCGTGAAGCAGCTCGTTCAGAAATCCAAGTATTAGAG 480
Db	492 ATCGTAAAAAATGTAGGCGGTTACCGTGAAGCAGCTCGTTCAGAAATCCAAGTATTAGAG 551
Qy	481 CACTTAAATAGTACTGATCCCAATAGTGTCTCCGATGTCTCCAGATGCCAGATGCTGAAGGTTT 540
Db	552 CACTTAAATAGTACTGATCCCAATAGTGTCTCTCCGATGTCTCCAGATGCTGAAGTGGTTT 611
Qy	541 GATCATCATGTGCATGTTTGTATTCTGTTTGAACCTACTGGGACTTAGTACTTACCATTTC 600
Db	612 GATCATCATGTGCATGTTTGTATTCTGTTTGAACCTACTGGGACTTAGTACTTACCATTTC 671
Qy	601 ATTAAGAAAAACAGCTTCTCGCAATTCAAATTGACCAATCAGGCGAGATGCGGTATCAG 660
Db	672 ATTAAGAAAAACAGCTTCTCGCAATTCAAATTGACCAATCAGGCGAGATGCGGTATCAG 731
Qy	661 ATCTGCCAGTCAATAAATTTTTTACATCATATAAAATTAACCCATACAGATCTGAAGGCT 720
Db	732 ATCTGCCAGTCAATAAATTTTTTACATCATATAAATTAACCCATACAGATCTGAAGGCT 791
Qy	721 GAAAAATTTTTGTTGTAAGTCGACACTATGTAGTCAAAATATAATCTTAAGTGAACCGT 780
Db	792 GAAAAATTTTTGTTGTAAGTCGACACTATGTAGTCAAAATATAATCTTAAGTGAACCGT 851
Qy	781 GATGAACGCACACTGAAAAACACAGATATCAAAGTTGTTGACTTTGGAAAGTCGAACGTAT 840
Db	852 GATGAACGCACACTGAAAAACACAGATATCAAAGTTGTTGACTTTGGAAAGTCGAACGTAT 911
Qy	841 GATGATGAATCATCAGTACTTTTGGTGTCTACCCGGCACTACAGAGCTCCCGAGGTCAATT 900


```
QY 1501 AATTTTCTTTTGAATTAACATTTTAAATACCTTTGAAATATCCTTTGATGTCACAGT 1560
Db 1572 AATTTTCTTTTGAATTAACATTTTAAATACCTTTGAAATATCCTTTGATGTCACAGT 1631
QY 1561 ATAAATGATGATGATCTTGGCTTTTGTACATGAGGTCACCTCTGAAGTGATTTTTTTTG 1620
Db 1632 ATAAATGATGATGATCTTGGCTTTTGTACATGAGGTCACCTCTGAAGTGATTTTTTTTG 1691
QY 1621 AGTAAAGGAAATCTTGACTACTTTATATCTTAAAGGAATATCTTTATATACCTTCAAA 1680
Db 1692 AGTAAAGGAAATCTTGACTACTTTATATCTTAAAGGAATATCTTTATATACCTTCAAA 1751
QY 1681 TTTAGAACCTTAACCTTTAAAGTTTTCTTCTGTAATTTGTTGAAACGGGTGATTTATTTAA 1740
Db 1752 TTTAGAACCTTAACCTTTAAAGTTTTCTTCTGTAATTTGTTGAAACGGGTGATTTATTTAA 1811
QY 1741 CTCTAGATAAGCAGTACTAGAACCAAACTCAGAAATCTTACTGTTAGATTTCTAT 1800
Db 1812 CTCTAGATAAGCAGTACTAGAACCAAACTCAGAAATCTTACTGTTAGATTTCTAT 1871
QY 1801 TAAATTTTAAAGTTTGTATTTCTTTTCTTGGTGATGTCAGGGTGATAACAGACATTC 1860
Db 1872 TAAATTTTAAAGTTTGTATTTCTTTTCTTGGTGATGTCAGGGTGATAACAGACATTC 1931
QY 1861 ATGAAAGGCAATGCAAGTTTGTCCATTTGACAGTTTGTATTTTAAATAAACAATACACT 1920
Db 1932 ATGAAAGGCAATGCAAGTTTGTCCATTTGACAGTTTGTATTTTAAATAAACAATACACT 1991
QY 1921 TTATTTAGATTTAAATCTTAAGTCTGAAAGTCAAGTCTGAAATGAGACATTTTCAAGTATG 1980
Db 1992 TTATTTAGATTTAAATCTTAAGTCTGAAAGTCAAGTCTGAAATGAGACATTTTCAAGTATG 2051
QY 1981 TTTGTGAGTCAAGATATAAATAAGAAATTTCTGATGAGAGTTTCAAGTTTAAATACC 2040
Db 2052 TTTGTGAGTCAAGATATAAATAAGAAATTTCTGATGAGAGTTTCAAGTTTAAATACC 2111
QY 2041 AAGTCCTTGGAGTCTTAACATTTGCCAGCATCTGTTTATCAAAATGACATAAATACGTA 2100
Db 2112 AAGTCCTTGGAGTCTTAACATTTGCCAGCATCTGTTTATCAAAATGACATAAATACGTA 2171
QY 2101 ACCTATAGATTAAGTTTATTAATAGGCAATTTATGTCGTGATTAATTTCTTACGGAG 2160
Db 2172 ACCTATAGATTAAGTTTATTAATAGGCAATTTATGTCGTGATTAATTTCTTACGGAG 2231
QY 2161 AAAGAGGATTTGATTTGAAAGCAGTTTCCGGAAGAAAGTCTGCTGAAATTTCCAGAAATTT 2220
Db 2232 AAAGAGGATTTGATTTGAAAGCAGTTTCCGGAAGAAAGTCTGCTGAAATTTCCAGAAATTT 2291
QY 2221 AATTGATTTGTTACATAAATCTTTTGACTTCAGAAA 2256
Db 2292 AATTGATTTGTTACATAAATCTTTTGACTTCAGAAA 2327

RESULT 4
AX392827 LOCUS 2354 bp DNA linear PAT 23-MAR-2002
DEFINITION Sequence 1 from Patent WO0216567.
ACCESSION AX392827
VERSION AX392827.1 GI:19700923
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yan, C.C., Ye, J.C., Ketchum, K.A., di Francesco, V.C. and Beasley, E.M.
TITLE Isolated human kinase proteins, nucleic acid molecules encoding
human kinase proteins, and uses thereof
JOURNAL Patent: WO 0216567-A 1 28-FEB-2002;
Applera Corporation (US)
FEATURES
Location/Qualifiers
source 1..2354
```

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

```
Query Match 100.0%; Score 2256; DB 6; Length 2354;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GACACTCATCCAGTCATTATTTAGAACCAAGGTCCTTTGAATGAGCGAGATTATCGGAC 60  
Db 72 GACACTCATCCAGTCATTATTTAGAACCAAGGTCCTTTGAATGAGCGAGATTATCGGAC 131  
QY 61 CGAGATACGTTGACGAATACAGGAATGACTACTGTGGAAGGATATGTTCTTAGACATTAT 120  
Db 132 CGAGATACGTTGACGAATACAGGAATGACTACTGTGGAAGGATATGTTCTTAGACATTAT 191  
QY 121 CACAGAGACATTGAAAGCGGGTATCGAATCCACTGCAGTAAATCTTCAGTCGGACGAG 180  
Db 192 CACAGAGACATTGAAAGCGGGTATCGAATCCACTGCAGTAAATCTTCAGTCGGACGAG 251  
QY 181 AGAAGCAGTCCTAAAGAGGAGCGCAATAGACACTGTTCAAGTCATCAGTCACGTTCCGAG 240  
Db 252 AGAAGCAGTCCTAAAGAGGAGCGCAATAGACACTGTTCAAGTCATCAGTCACGTTCCGAG 311  
QY 241 AGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGT 300  
Db 312 AGCCACCGAAGAAAGATCCAGGAGTATAGAGGATGATGAGGAGGTCACCTGATCTGT 371  
QY 301 CAAGTGGAGACGTTCTAAGAGCAAGATATCAAAATCGTGGACACTTTGGGTGAAGGAGCC 360  
Db 372 CAAGTGGAGACGTTCTAAGAGCAAGATATCAAAATCGTGGACACTTTGGGTGAAGGAGCC 431  
QY 361 TTGCGAAAAGTTGTAGAGTGCAATTCATCGCATGATGATGATGATGATGATGATGATGAT 420  
Db 432 TTGCGAAAAGTTGTAGAGTGCAATTCATCGCATGATGATGATGATGATGATGATGATGAT 491  
QY 421 ATCGTAAAAATGTAGGCGGTTACCGTGAAGCAGCTCGTTCAGAAAATCCAAGTATTAGAG 480  
Db 492 ATCGTAAAAATGTAGGCGGTTACCGTGAAGCAGCTCGTTCAGAAAATCCAAGTATTAGAG 551  
QY 481 CACTTAATAGTACTGATCCCAATAGTGTCTTCCGATGTGTCAGATGCTAGATGCTTT 540  
Db 552 CACTTAATAGTACTGATCCCAATAGTGTCTTCCGATGTGTCAGATGCTAGATGCTTT 611  
QY 541 GATCATCATGTCATGTTTGTATTGTTTCAACTACTGCGACCTTAGTACTTACGATTTTC 600  
Db 612 GATCATCATGTCATGTTTGTATTGTTTCAACTACTGCGACCTTAGTACTTACGATTTTC 671  
QY 601 ATTAAGAAAAACAGCTTTTCCCAATTTCAAAATTGACCACATCAGGACAGATGCGGTATCAG 660  
Db 672 ATTAAGAAAAACAGCTTTTCCCAATTTCAAAATTGACCACATCAGGACAGATGCGGTATCAG 731  
QY 661 ATCTGCCAGTCATATAAATTTTTCATCATATAAATTAACCCATACAGATCTGAAGCTT 720  
Db 732 ATCTGCCAGTCATATAAATTTTTCATCATATAAATTAACCCATACAGATCTGAAGCTT 791  
QY 721 GAAAAATATTTTGTGTAAGTCTGACTATGCTAGTCAAAATATAATTTTAAATGAAACGT 780  
Db 792 GAAAAATATTTTGTGTAAGTCTGACTATGCTAGTCAAAATATAATTTTAAATGAAACGT 851  
QY 781 GATGAACGCACACTGAAAAACACAGATATCAAAAGTTGTTGACTTTGGAAGTGCACAGTAT 840  
Db 852 GATGAACGCACACTGAAAAACACAGATATCAAAAGTTGTTGACTTTGGAAGTGCACAGTAT 911  
QY 841 GATGATGAACATACAGTACTTTGGTGTCTACCGGCACTACAGAGCTCCGAGGTCAAT 900  
Db 912 GATGATGAACATACAGTACTTTGGTGTCTACCGGCACTACAGAGCTCCGAGGTCAAT 971  
QY 901 TTGGCTTTAGTTGGTCTCAGCCCTTGATGCTTTGGAGCATAGGTTGCTATTTATTCGA 960  
Db 972 TTGGCTTTAGTTGGTCTCAGCCCTTGATGCTTTGGAGCATAGGTTGCTATTTATTCGA 1031
```

Qy	961	TATTTACCTTGGTTTTCACAGCTCTTTACAGCTCATGATAGTAAGAGCACCTGGCAATGATG	1021
Db	1032	TATTTACCTTGGTTTTCACAGCTCTTTACAGCTCATGATAGTAAGAGCACCTGGCAATGATG	1091
Qy	1021	GAAACGAATATTTAGGACCCATACCAACAACATGATTCAGAAAAACAGAAAAACCAAGTAT	1080
Db	1092	GAAACGAATATTTAGGACCCATACCAACAACATGATTCAGAAAAACAGAAAAACCAAGTAT	1151
Qy	1081	TTTTCACCATTAACCACTAGATTGGGATGAACACAGTTCTGCTGGTAGATATGTTAGGAGA	1140
Db	1152	TTTTCACCATTAACCACTAGATTGGGATGAACACAGTTCTGCTGGTAGATATGTTAGGAGA	1211
Qy	1141	CGCTGCAAAACCGTTGAAAGGAATTTATGCTTTGTGCATGATGAAGAAACATGAGAAACTGTTT	1200
Db	1212	CGCTGCAAAACCGTTGAAAGGAATTTATGCTTTGTGCATGATGAAGAAACATGAGAAACTGTTT	1271
Qy	1201	GACCTGGTTCCGAAGAAATGTTAGATATGATCCCACTCAAGAAATTAACCTTGGATGAAGCA	1260
Db	1272	GACCTGGTTCCGAAGAAATGTTAGATATGATCCCACTCAAGAAATTAACCTTGGATGAAGCA	1331
Qy	1261	TTGCAGCATCCCTTTCTTTGACTTATTAATAAAGAAATGAATGGAAATTAACCTTGGATGAAGCA	1320
Db	1332	TTGCAGCATCCCTTTCTTTGACTTATTAATAAAGAAATGAATGGAAATTAACCTTGGATGAAGCA	1391
Qy	1321	TATATACCTCTCTAGAGAGATTACTTAAGACTGTGTCACTCAATCAATCTCAATAT	1380
Db	1392	TATATACCTCTCTAGAGAGATTACTTAAGACTGTGTCACTCAATCAATCTCAATAT	1451
Qy	1381	TTTTTGTAAACATTAATAATTTTGTACAGTTTAAGTGTAAATNTGTATGTTTGTATCAA	1440
Db	1452	TTTTTGTAAACATTAATAATTTTGTACAGTTTAAGTGTAAATNTGTATGTTTGTATCAA	1511
Qy	1441	TAGCATTAATTAACCTTGTAAACCAAGATGGTCTTGATAATGCATTAGAAAAATTAATAAT	1500
Db	1512	TAGCATTAATTAACCTTGTAAACCAAGATGGTCTTGATAATGCATTAGAAAAATTAATAAT	1571
Qy	1501	AAATTTTCTTTTGAATAACCAATTTTAAATACCTTTGAAAAATACCTTTGTGCGAGTG	1560
Db	1572	AAATTTTCTTTTGAATAACCAATTTTAAATACCTTTGAAAAATACCTTTGTGCGAGTG	1631
Qy	1561	ATAAATGTGATTTGATCTTGCCCTTTTGATCATGGAGGTCACTCTGAAGTCAATTTTTTTTG	1620
Db	1632	ATAAATGTGATTTGATCTTGCCCTTTTGATCATGGAGGTCACTCTGAAGTCAATTTTTTTTG	1691
Qy	1621	AGTAAAGGAAATCTTGACTACTTTATATCTTAAAGGAATATCTTTATATATCTTCAAA	1680
Db	1692	AGTAAAGGAAATCTTGACTACTTTATATCTTAAAGGAATATCTTTATATATCTTCAAA	1751
Qy	1681	TTTAGAACTTAACCTTTAAAGTTTTTCTTCTGTAATTTGTGAACGGGTGATTAATATAA	1740
Db	1752	TTTAGAACTTAACCTTTAAAGTTTTTCTTCTGTAATTTGTGAACGGGTGATTAATATAA	1811
Qy	1741	CTCTAGATAAGCAGGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGAAATCTAT	1800
Db	1812	CTCTAGATAAGCAGGTACTAGAAACCAAACTCAGAAAAATGTTTACTGTTAGAAATCTAT	1871
Qy	1801	TAAATTTTAAAGTGTGTATCTTTTCTTCAATGGGTGATGTCAGGGTGATTAACCAAGATTC	1860
Db	1872	TAAATTTTAAAGTGTGTATCTTTTCTTCAATGGGTGATGTCAGGGTGATTAACCAAGATTC	1931
Qy	1861	ATGGAAGGCAATGCAAGTTTGTCCATTTGTGACAGTTTGTGTAATAAAAAACCAACACCT	1920
Db	1932	ATGGAAGGCAATGCAAGTTTGTCCATTTGTGACAGTTTGTGTAATAAAAAACCAACACCT	1991
Qy	1921	TTATTTAAGATTAATACTTAACGTGGAAGTCACTTGGAAAAATGGGAACATTTCCAGATGATG	1980
Db	1992	TTATTTAAGATTAATACTTAACGTGGAAGTCACTTGGAAAAATGGGAACATTTCCAGATGATG	2051
Qy	1981	TTTTCGTGAGTCCACAGATATAAAATAGAAATCTTGATCAGAGGTTTCAGTTTTTAAATACC	2040
Db	2052	TTTTCGTGAGTCCACAGATATAAAATAGAAATCTTGATCAGAGGTTTCAGTTTTTAAATACC	2111
Qy	2041	AAGTCTTTAGGAGTCTTAAACATTTGGCCACGACCTCTGTTTATCAAAATGACATTAATAACGTAA	2100

Db	2112	AAAGTCCTTAGGAGTCTTAACATTTGGCAGCACTGTTTATCAANTGCAATAATACGTAA	2171
Qy	2101	ACCTATAAGCAATTAAGTTTATTAAATTAGGCAATTTATGCTGTGATAAATCTTTACGGGAG	2160
Db	2172	ACCTATAAGCAATTAAGTTTATTAAATTAGGCAATTTATGCTGTGATAAATCTTTACGGGAG	2231
Qy	2161	AAAGAGATTTGATTTGGAAGACGATTTGGGAAGAAAGTCTGCTGAAATTTCCAGAAATTT	2220
Db	2232	AAAGAGATTTGATTTGGAAGACGATTTGGGAAGAAAGTCTGCTGAAATTTCCAGAAATTT	2291
Qy	2221	AAATGATTTGTTACATAAACTTTTTCGACTTCAGAAA	2256
Db	2292	AAATGATTTGTTACATAAACTTTTTCGACTTCAGAAA	2327
RESULT 5			
AX056410			
LOCUS	AX056410	2488 bp	DNA linear PAT 13-JAN-2001
DEFINITION	Sequence 54 from Patent WO0073469.		
ACCESSION	AX056410		
VERSION	AX056410.1	GI:12229117	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 Ploman G.D., Martinez, R., Whyte, D. and Sudersanam, S.		
TITLE	Protein kinases		
JOURNAL	Patent: WO 0073469-A 54 07-DEC-2000;		
FEATURES	Sugen, Inc. (US)		
source	Location/Qualifiers		
	1..2488		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	90.8%;	Score 2048;	DB 6; Length 2488;
Best Local Similarity	95.9%;	Pred. No. 0;	
Matches 2148; Conservative	0;	Mismatches 1;	Indels 90; Gaps 1;
Qy	14	GTCAATTATTAGAACGAAGTCTCTGAATGAGCGAGATTATCGGGACCGGAGATACGTTG	73
Db	338	GTCAATTATTAGAACGAAGTCTCTGAATGAGCGAGATTATCGGGACCGGAGATACGTTG	397
Qy	74	ACGAATACAGGAATGACTACTGTGGAAGGATATGTTCTTACACATATTCACAGAGACATTG	133
Db	398	ACGAATACAGGAATGACTACTGTGGAAGGATATGTTCTTACACATATTCACAGAGACATTG	457
Qy	134	AAAGCGGGTATCGAATCCAATCGACATTAATCTTTCAAGTCCGACGAGGAAGCAGTCTTA	193
Db	458	AAAGCGGGTATCGAATCCAATCGACATTAATCTTTCAAGTCCGACGAGGAAGCAGTCTTA	517
Qy	194	AAAGGAAGCGCAATAGACACTGTTTCAAGTCAATCAGTCACTGTTGGAAGAGCCACCGAAGGA	253
Db	518	AAAGGAAGCGCAATAGACACTGTTTCAAGTCAATCAGTCACTGTTGGAAGAGCCACCGAAGGA	561
Qy	254	AAAGATCCAGGAGTATAGAGATGATGAGAGGGTCACTGATCTGTCAAAGTGGAGACG	313
Db	562	-----	561
Qy	314	TTCTAAGACGAAGATATGAATCTGTGACACTTTGGGTGAAGAGCCTTTGGCAAAAGTTG	373
Db	562	-----NATGAATCTGTGACACTTTGGGTGAAGAGCCTTTGGCAAAAGTTG	607
Qy	374	TAGAGTGCATTTGATCATGGCATGGATGGCATGATGTAGCAGTCAAAATCGTAAAAAATG	433
Db	608	TAGAGTGCATTTGATCATGGCATGGATGGCATGATGTAGCAGTCAAAATCGTAAAAAATG	667
Qy	434	TAGCGCGTTACCGTGAAGCAGCTCGTTCAAGATCCAAGTATTAGACACTTAATAGTA	493

Db 668 TAGGCGGTACCGTGAAGAGAGCTCGTTCCAGAAATCCAAAGTATTAGAGCAGCTTAAATAGTA 727
Qy 494 CTGATCCCAATAGTGTCTTCGATGTGTCAGATGCTAGAAATGGTTTGGATCATCATGTC 553
Db 728 CTGATCCCAATAGTGTCTTCGATGTGTCAGATGCTAGAAATGGTTTGGATCATCATGTC 787
Qy 554 ATGTTTGTATGTTGTTGAACTACTGCGGACCTAGTACTTACGATTTCAATTAAGAAACA 613
Db 788 ATGTTTGTATGTTGTTGAACTACTGCGGACCTAGTACTTACGATTTCAATTAAGAAACA 847
Qy 614 GCTTTCTGCAATTCAGAAATGACACATCAGGACATGCGGTATCAGATCTGCCAGTCAA 673
Db 848 GCTTTCTGCAATTCAGAAATGACACATCAGGACATGCGGTATCAGATCTGCCAGTCAA 907
Qy 674 TAAATTTTTTACATCATATAATTAATTAACCCATACAGATCTGAAGCCTGAAATATTTTGT 733
Db 908 TAAATTTTTTACATCATATAATTAATTAACCCATACAGATCTGAAGCCTGAAATATTTTGT 967
Qy 734 TTGTGAAGTCTGATGATGATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 793
Db 968 TTGTGAAGTCTGATGATGATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1027
Qy 794 TGAATAACACAGATATCAAGTGTGTTGACTTTTGAAGTGCACGATGATGATGAACATC 853
Db 1028 TGAATAACACAGATATCAAGTGTGTTGACTTTTGAAGTGCACGATGATGATGAACATC 1087
Qy 854 ACAGTACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 913
Db 1088 ACAGTACTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1147
Qy 914 GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCAATCTTATTAATTAATTAATTAATTAAT 973
Db 1148 GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGCAATCTTATTAATTAATTAATTAATTAAT 1207
Qy 974 TCACAGTCTTTCAGTCTCATGATGATGAAGCAGCTGCGCATGATGGAAGATATTAG 1033
Db 1208 TCACAGTCTTTCAGTCTCATGATGATGAAGCAGCTGCGCATGATGGAAGATATTAG 1267
Qy 1034 GACCCATACCAACACACATGATTTCAGAAAAACGAAAGATATTTCACCATATACC 1093
Db 1268 GACCCATACCAACACACATGATTTCAGAAAAACGAAAGATATTTCACCATATACC 1327
Qy 1094 AGCTAGATTGGATGAACACAGTCTGCTGGTAGATATGTTAGGAGAGCGTGCAACCGT 1153
Db 1328 AGCTAGATTGGATGAACACAGTCTGCTGGTAGATATGTTAGGAGAGCGTGCAACCGT 1387
Qy 1154 TGAAGGAATTTATGCTTGTCTGATGATGAAGACATGAGAACTGTTGACCTGGTTCGAA 1213
Db 1388 TGAAGGAATTTATGCTTGTCTGATGATGAAGACATGAGAACTGTTGACCTGGTTCGAA 1447
Qy 1214 GAATGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGAGCATCCTT 1273
Db 1448 GAATGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGAGCATCCTT 1507
Qy 1274 TCTTTGACTTATTAATAAAGAAATGAAATGGAAATCAGTGGTCTTATATATCTCTCT 1333
Db 1508 TCTTTGACTTATTAATAAAGAAATGAAATGGAAATCAGTGGTCTTATATATCTCTCT 1567
Qy 1334 AGAAGAGATTACTTAAGTGTGTCAGTCAACTAAACATTTCTAATTTTGTAAACATT 1393
Db 1568 AGAAGAGATTACTTAAGTGTGTCAGTCAACTAAACATTTCTAATTTTGTAAACATT 1627
Qy 1394 AAAATTTTGTGACAGTTAAAGTGAATATTTGTTTGTATCAATGACATATTAAC 1453
Db 1628 AAAATTTTGTGACAGTTAAAGTGAATATTTGTTTGTATCAATGACATATTAAC 1687
Qy 1454 TTGTTAAGCAAGTATGTTGATTAATGATTAAGAAAAATTAATAATTTTCTTTT 1513
Db 1688 TTGTTAAGCAAGTATGTTGATTAATGATTAAGAAAAATTAATAATTTTCTTTT 1747
Qy 1514 GAAATACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGAATGATGATG 1573
Db 1748 GAAATACCAATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGAATGATGATG 1807

Qy 1574 ATCTTGCCCTTTGTACATGAGGTCACCTCTGAAGTGAATTTTTTTTGTAGTAAAGGAAAT 1633
Db 1808 ATCTTGCCCTTTGTACATGAGGTCACCTCTGAAGTGAATTTTTTTTGTAGTAAAGGAAAT 1867
Qy 1634 CTTGACTACTTTATATTTCTTAAAGGAATATTTCTTTATATATCTCAAAATTTAGAACTTAAC 1693
Db 1868 CTTGACTACTTTATATTTCTTAAAGGAATATTTCTTTATATATCTCAAAATTTAGAACTTAAC 1927
Qy 1694 TTTAAAGTTTTTCTTCTGTAATTTGTAACGGGTGATTTATTTAACTCTAGATAAGCA 1753
Db 1928 TTTAAAGTTTTTCTTCTGTAATTTGTAACGGGTGATTTATTTAACTCTAGATAAGCA 1987
Qy 1754 GGTACTAGAAACCAAACTCAGAAAAATGTTTCTGTTTGAATTTCTATTAATTTTAAAGTG 1813
Db 1988 GGTACTAGAAACCAAACTCAGAAAAATGTTTCTGTTTGAATTTCTATTAATTTTAAAGTG 2047
Qy 1814 TTGTAATTTTTTCTTCTGTAATTTGTAACGGGTGATTTATTTAACTCTAGATAAGCA 1873
Db 2048 TTGTAATTTTTTCTTCTGTAATTTGTAACGGGTGATTTATTTAACTCTAGATAAGCA 2107
Qy 1874 CAGTTTGTCCATTTGTGACAGTTTGTAAATAAAACCAATACACACTTTTATTTAAAGTAA 1933
Db 2108 CAGTTTGTCCATTTGTGACAGTTTGTAAATAAAACCAATACACACTTTTATTTAAAGTAA 2167
Qy 1934 AAATCTAACTGGAAGTCTGAGTTTGTGAAAAATGCAATTTTCCAAAGTATGTTTGTGAGTCA 1993
Db 2168 AAATCTAACTGGAAGTCTGAGTTTGTGAAAAATGCAATTTTCCAAAGTATGTTTGTGAGTCA 2227
Qy 1994 AGATATAAAATAGAAATTTCTGATGAGAGGTTTCAAGTTTAAATACCAAGTCTTTAGGAG 2053
Db 2228 AGATATAAAATAGAAATTTCTGATGAGAGGTTTCAAGTTTAAATACCAAGTCTTTAGGAG 2287
Qy 2054 TCTTAAACATTCGCCAGCATCTGTTTATCAAAATGACATAAAATACGTAAACCTTATAGAAT 2113
Db 2288 TCTTAAACATTCGCCAGCATCTGTTTATCAAAATGACATAAAATACGTAAACCTTATAGAAT 2347
Qy 2114 AAGTTTATTAATAGGCAATTTATGTTCTGTGATAATTTCTTACGGGAGAAAGAGGATTTGA 2173
Db 2348 AAGTTTATTAATAGGCAATTTATGTTCTGTGATAATTTCTTACGGGAGAAAGAGGATTTGA 2407
Qy 2174 TTGAAAGCAGTTTGGGAGAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2233
Db 2408 TTGAAAGCAGTTTGGGAGAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2467
Qy 2234 CATAACTTTTGTGCTTCA 2252
Db 2468 CATAACTTTTGTGCTTCA 2486

RESULT 6

AR077239 LOCUS AR077239 2446 bp DNA linear PAT 31-AUG-2000
DEFINITION Sequence 9 from patent US 5962232.

AR077239 ACCESSION AR077239

AR077239.1 GI:10003985

KEYWORDS

SOURCE Unknown.

ORGANISM

Unclassified.

1 (bases 1 to 2446)

Bandman,O., Hillman,J.L., Lal,P., Akerblom,I.E., Shah,P.,

Corley,N.C. and Guegler,K.J.

Protein kinase molecules

Patent: US 5962232-A 9 05-OCT-1999;

Location/Qualifiers

1..2446

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Query Match 90.7%; Score 2046.6; DB 6; Length 2446;

Best Local Similarity 95.8%; Pred. No. 0;

[illegible]

QY	1394	AAATTAATTTGACAGTTAAAGTGAATTAATTTGATGTTTGTATCAATAGCATAATTAAC	1453
Db	1578	AAATTAATTTTGTACAGTTAAAGTGAATTAATTTGATGTTTGTATCAATAGCATAATTAAC	1637
QY	1454	TGTTAAGCAAGTATGGTCTTGATTAATGATAGAAAATTAATTAATTTTCTTTT	1513
Db	1638	TGTTAAGCAAGTATGGTCTTGATTAATGATAGAAAATTAATTAATTTTCTTTT	1697
QY	1514	GAAATTAACATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATAAATGATG	1573
Db	1698	GAAATTAACATTTTAAATACCTTTGAAATATCCTTTGTGTCAGTGATAAATGATG	1757
QY	1574	ATCTTGCCCTTTTGATACATGAGGTCACCTCTGAAGTATTTTGTAGTAAAGGAAT	1633
Db	1758	ATCTTGCCCTTTTGATACATGAGGTCACCTCTGAAGTATTTTGTAGTAAAGGAAT	1817
QY	1634	CTTGACTACTTTTATATCTTAAAGGAATATCTTTTATATATCTTCAAAATTTAGAACTTAAC	1693
Db	1818	CTTGACTACTTTTATATCTTAAAGGAATATCTTTTATATATCTTCAAAATTTAGAACTTAAC	1877
QY	1694	TTTAAAGTTTTTCTCTGTAATTTGTTGAACGGGTGATTTATTTAACTCTAGATAAGCA	1753
Db	1878	TTTAAAGTTTTTCTCTGTAATTTGTTGAACGGGTGATTTATTTAACTCTAGATAAGCA	1937
QY	1754	GGTACTAGAAACCAAACTCAGAAAATGTTTACTGTTAGAAATCTTATTAATTTTAACTG	1813
Db	1938	GGTACTAGAAACCAAACTCAGAAAATGTTTACTGTTAGAAATCTTATTAATTTTAACTG	1997
QY	1814	TGTTATCTTTTTCATTTGGGTGATGTCAGGGTGATAACCGAGCATTCATGGAAGGCATG	1873
Db	1998	TGTTATCTTTTTCATTTGGGTGATGTCAGGGTGATAACCGAGCATTCATGGAAGGCATG	2057
QY	1874	CAGTTTGTCCATTGTGACAGTTTGTGTTAATAAACCACATACACACTTTTATTAAGATTA	1933
Db	2058	CAGTTTGTCCATTGTGACAGTTTGTGTTAATAAACCACATACACACTTTTATTAAGATTA	2117
QY	1934	AAATCTAACTGGAAGTACAGCTTGGAATAATGGACATTTCCAAGTATGTTGGTGAGTCAC	1993
Db	2118	AAATCTAACTGGAAGTACAGCTTGGAATAATGGACATTTCCAAGTATGTTGGTGAGTCAC	2177
QY	1994	AGATATAAATAAGAAATCTGATGAGAGTTTCAAGTTTAAATACCAAGTCTTTAGGAG	2053
Db	2178	AGATATAAATAAGAAATCTGATGAGAGTTTCAAGTTTAAATACCAAGTCTTTAGGAG	2237
QY	2054	TCTTAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAATACGTAACCTATAAGAAAT	2113
Db	2238	TCTTAACATTTGGCCAGCATCTGTTTATCAAAATGACATAAATACGTAACCTATAAGAAAT	2297
QY	2114	AAGTTTATTAATTAGGCAATTTATGTCGTGATAATCTTACGGGAGAAAGAGGATTTGA	2173
Db	2298	AAGTTTATTAATTAGGCAATTTATGTCGTGATAATCTTACGGGAGAAAGAGGATTTGA	2357
QY	2174	TTGGAAGCAGTTTGGGAAGAGAGTGTCTGAAATTTCCAAGAAATTAATGATTTGGTTA	2233
Db	2358	TTGGAAGCAGTTTGGGAAGAGAGTGTCTGAAATTTCCAAGAAATTAATGATTTGGTTA	2417
QY	2234	CATAAACTTTTGACTTCAGAAA	2256
Db	2418	CATAAACTTTTGACTTCAGAAA	2440
RESULT 8			
AF294429			
LOCUS	1814 bp mRNA linear PRI 01-MAR-2001		
DEFINITION	Homo sapiens protein serine threonine kinase Clk4 mRNA, complete cds.		
ACCESSION	AF294429		
VERSION	AF294429.1 GI:9965397		
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			

REFERENCE	1	(bases 1 to 1814)	Schultz, J., Jones, T., Bork, P., Sheer, D., Blencke, S., Steyrer, S., Wellbrock, U., Bevec, D., Ullrich, A. and Wallasch, C.
AUTHORS			Molecular characterization of a cDNA encoding functional human CLK4 kinase and localization to chromosome 5q35 [correction of 4q35]
TITLE			Genomics 71 (3), 368-370 (2001)
JOURNAL			21100912
MEDLINE			11170754
PUBMED			
REFERENCE	2	(bases 1 to 1814)	Schultz, J., Jones, T., Bork, P., Sheer, D., Blencke, S., Steyrer, S., Wellbrock, U., Bevec, D., Ullrich, A. and Wallasch, C.
AUTHORS			Direct Submission
TITLE			Submitted (08-AUG-2000) Platform Technology, Axxina Pharmaceuticals AG, Am Klopferspitz 19, Martinsried 82152, Germany
JOURNAL			Location/Qualifiers
FEATURES			1. 1814
source			/organism="Homo sapiens"
			/mol_type="mRNA"
			/db_xref="taxon:9606"
			/chromosome="4"
			/map="4q35"
			1. 1446
CDS			/codon_start=1
			/product="protein serine threonine kinase Clk4"
			/protein_id="AAG10074.1"
			/db_xref="GI:9965398"
			/translation="MRHSKRTHCPDWDSRESWGSHGSRKRRKRSHTSQENRHCK PHQFKSDCHYLEARSINERDRYRDYCEGVYPRVHRDIESGVRIHCS KSSVRSRSPKRNHRKSSHOSRSHRKRKRSIEDDEGHLICQSGDLVLRVY IVDTLGSGAGKGVCEIDHGMHVAKIVNRYREARSEIOVLHNLNTPNS VFRQVMEWFDDHGHVCI VFELGLSTLDFIKENSFLPFQIDHROMAYQICOSINF LHNKLTHLDPENILFVKSQYVYVKNRMDERTLKNTDIKVDFGSGATYDEHH STLVSTRHYRAPEVILALGWSQPCDVMSIGCLIEYLGFTVFTQTHDSKEHLAMERI LGPIQHMIOKTRKRKVFHNLQDWEHSSAGRYVRRRCKPLKEFMLCHDEHEKLFID LVREMLEVDPTQRTLDLALQHPFFDLKKK"
ORIGIN			
Query Match	72.3%;	Score 1632;	DB 9; Length 1814;
Best Local Similarity	99.3%;	Pred. No. 1.3e-292;	
Matches 1638;	Conservative 0;	Mismatches 11;	Indels 0; Gaps 0;
QY	14	GTCAATATTAGAACGAGTCTTGAATGACGAGATATCGGACCGGAGATACGTTG	73
Db	161	GTCAATATTAGAACGAGTCTTGAATGACGAGATATCGGACCGGAGATACGTTG	220
QY	74	ACGAATACAGGAATGACTATCTGTAAGGATATGTTCTTAGACATTATCACAGACATG	133
Db	221	ACGAATACAGGAATGACTATCTGTAAGGATATGTTCTTAGACATTATCACAGACATG	280
QY	134	AAAGCGGTATCGAATCCATCGAGTAAATCTTCAGTCCGACGAGGAGAGCAGTCCTA	193
Db	281	AAAGCGGTATCGAATCCATCGAGTAAATCTTCAGTCCGACGAGGAGAGCAGTCCTA	340
QY	194	AAAGGAACGCAATAGACACCTGTTCACTCATCTAGTACGTTTGAAGAGCACCAGGA	253
Db	341	AAAGGAACGCAATAGACACCTGTTCACTCATCTAGTACGTTTGAAGAGCACCAGGA	400
QY	254	AAAGATCCAGGATATAGAGGATGATAGAGGGTCACTCATCTGTCTGTAAGTGGAG	313
Db	401	AAAGATCCAGGATATAGAGGATGATAGAGGGTCACTCATCTGTCTGTAAGTGGAG	460
QY	314	TTCTAAGAGCAAGATATGAAATCGTGGACACATTTGGGTGAAGAGCCTTTGGCAAG	373
Db	461	TTCTAAGAGCAAGATATGAAATCGTGGACACATTTGGGTGAAGAGCCTTTGGCAAG	520
QY	374	TAGAGTGCATTCATCGCATGCGCATGTCATGTCATGTCATGTCATGTCATGTCATG	433
Db	521	TAGAGTGCATTCATCGCATGCGCATGTCATGTCATGTCATGTCATGTCATGTCATG	580
QY	434	TAGGCGGTTACCGTGAAGCAGCTCGTTCCAGAAATCCAAAGTATTAGAGCATTAAAT	493
Db	581	TAGGCGGTTACCGTGAAGCAGCTCGTTCCAGAAATCCAAAGTATTAGAGCATTAAAT	640

```
QY 494 CTGATCCCAATAGTGTCTTCGATGTGTCAGATGCTAGATGGTTGATCATCATGTC 553
Db 641 CTGATCCCAATAGTGTCTTCGATGTGTCAGATGCTAGATGGTTGATCATCATGTC 700
QY 554 ATGTTTGTATGCTGTTTGAACCTACTGGGACTTAGTACTTACGATTTTCAATTAAGAAACA 613
Db 701 ATGTTTGTATGCTGTTTGAACCTACTGGGACTTAGTACTTACGATTTTCAATTAAGAAACA 760
QY 614 GCTTTCTGCCATTTCAAAATGACCATCATCAGGCAGATGGCGTATCAGATCTGCCAGTCAA 673
Db 761 GCTTTCTGCCATTTCAAAATGACCATCATCAGGCAGATGGCGTATCAGATCTGCCAGTCAA 820
QY 674 TAAATTTTTTACATCATATAAATTAACCCATACAGATCTGAAGCTGAATATATTTGT 733
Db 821 TAAATTTTTTACATCATATAAATTAACCCATACAGATCTGAAGCTGAATATATTTGT 880
QY 734 TTGTGAAGTCTGACATGATGAGTCAAAATATAATTTCTAAATGAAGCTGATGAAGCCACAC 793
Db 881 TTGTGAAGTCTGACATGATGAGTCAAAATATAATTTCTAAATGAAGCTGATGAAGCCACAC 940
QY 794 TGAATAAACACAGATATCAAAAGTTGTTGACTTTTGGAAAGTGCACAGTATGATGAACATC 853
Db 941 TGAATAAACACAGATATCAAAAGTTGTTGACTTTTGGAAAGTGCACAGTATGATGAACATC 1000
QY 854 ACAGTACTTTGTTGCTACCCGGCACTACAGAGTCCCGAGTCAATTTTGGCTTTAGGTT 913
Db 1001 ACAGTACTTTGTTGCTACCCGGCACTACAGAGTCCCGAGTCAATTTTGGCTTTAGGTT 1060
QY 914 GGTCTCAGCTTGTCATGTTTGGAGCATAGGTTGATTTCTTTATGAATATACCTTGGTT 973
Db 1061 GGTCTCAGCTTGTCATGTTTGGAGCATAGGTTGATTTCTTTATGAATATACCTTGGTT 1120
QY 974 TCACAGTCTTTTCAGACTCATGATAGTAAAGACACCTGGCAATGATGGAACGAATATTAG 1033
Db 1121 TCACAGTCTTTTCAGACTCATGATAGTAAAGACACCTGGCAATGATGGAACGAATATTAG 1180
QY 1034 GACCCATACCAACACATGATTCAGAAAACGAAGAAACGCAAGTATTTTACCACTAAC 1093
Db 1181 GACCCATACCAACACATGATTCAGAAAACGAAGAAACGCAAGTATTTTACCACTAAC 1240
QY 1094 AGCTAGATGGGATGAACACAGTCTCTGCTGGTAGATATGTTAGGAGAGCGTGCACACCGT 1153
Db 1241 AGCTAGATGGGATGAACACAGTCTCTGCTGGTAGATATGTTAGGAGAGCGTGCACACCGT 1300
QY 1154 TGAAGGAATTTATGCTTTGTCATGATGAAGAACATGAGAAAATGTTTGCACCTGGTTGAA 1213
Db 1301 TGAAGGAATTTATGCTTTGTCATGATGAAGAACATGAGAAAATGTTTGCACCTGGTTGAA 1360
QY 1214 GAATGTTAGATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCATTTGCAGCATCCTT 1273
Db 1361 GAATGTTAGATATGATCCAACTCAAGAAATTAACCTTGGATGAAGCATTTGCAGCATCCTT 1420
QY 1274 TCTTTGACTTATTAAGAAAGAAATGAAATGGAATCAGTGGTCTTACTATATCTCT 1333
Db 1421 TCTTTGACTTATTAAGAAAGAAATGAAATGGAATCAGTGGTCTTACTATATCTCT 1480
QY 1334 AGAAGAGATTAATTAAGACTGTGTCACTCAACCTAAACATTTCTAATATTTTGTAAACAT 1393
Db 1481 AGAAGAGATTAATTAAGACTGTGTCACTCAACCTAAACATTTCTAATATTTTGTAAACAT 1540
QY 1394 AAATTTATTTGATGAGTTAAATGTAATTTGATGTTTGTGTTGATCAATAGCATATAATAC 1453
Db 1541 AAATTTATTTGATGAGTTAAATGTAATTTGATGTTTGTGTTGATCAATAGCATATAATAC 1600
QY 1454 TTGTTAAGCAAGTATGGTCTTGAATATGATGATGATGATGATGATGATGATGATGATGAT 1513
Db 1601 TTGTTAAGCAAGTATGGTCTTGAATATGATGATGATGATGATGATGATGATGATGATGAT 1660
QY 1514 GAAATTTACCATTTTAAATATACCTTTGAAATATCTTTGTTGTCAGTGAATGATGATGAT 1573
Db 1661 GAAATTTACCATTTTAAATATACCTTTGAAATATCTTTGTTGTCAGTGAATGATGATGAT 1720
QY 1574 ATCTTGGCTTTTGTATGATGGAGGTCACTCTCTGAAGTGAATTTTTTTTTTGTAGTAAAGGAAAT 1633
```

```
Db 1721 ATCTTGCCCTTTGTACATGGAGGTCACTCTGAAGTGAATTTTTTGTAGTAAAGGAAAT 1780
QY 1634 CTGACTACTTTTATATTTCTTAAAGGAATA 1662
Db 1781 CTGACTACTTTTAAAGGAAATTTTTTGTAGTAAAGGAAAT 1809

RESULT 9
AX961896 1792 bp DNA linear PAT 14-JAN-2004
LOCUS Sequence 107 from Patent WO03104277.
DEFINITION AX961896
ACCESSION AX961896
VERSION AX961896.1 GI:40881306
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Sugahara,T., Mateuda,A., Honda,G., Muramatsu,S. and Ishizawa,K.
TITLE Stat6 activation gene
JOURNAL Patent: WO 03104277-A 107 18-DEC-2003;
Asahi Kasei Kabushiki Kaisha (JP)
FEATURES
source
1. .1792
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1. .1446
/note="unnamed protein product"
/codon_start=1
/protein_id="CAF06457.1"
/db_xref="GI:40881307"
/translation="MRHSKRTHCPDMSRSGHESYRSHKRRSHSSSTOENPHCK
PHQKSDCHYLEARSLNERDYDRRYDYRNDYCEGYVPRHYRHRDIESGRYHCS
KSSVRSRRSPRRKRRSHSRSHRRSRSDIEDDEGHLICOGDVLRYRYS
IVDTLGEAGFKGVKEIDHGMGHVAKIVKNGVREARASIVLEHLNSIDPNS
VFRQVQLEWDFHGHVCIIVFELGLSTYDFIKENSLFQIDHIRMAYQICQINF
LHNKLTHTDLKPNILFKVSDYVKNKMRDERTLKNTDIKVDFGSATYDDEHH
SLTVSTRHYRAPEVILALGWSQPCDMSIGCILLEYLGVTFVOTHSKEHLAMMERI
LGPIDQMIQKTRKRYFHNQLDWDHSSSAGRYVRRCKPLKPEMLCHDEHEKFLD
LVRRNLSYDTQRTITLDEALQHPFFDLKKK"

ORIGIN
Query Match 72.2%; Score 1629.4; DB 6; Length 1792;
Best Local Similarity 99.9%; Pred. No. 3.9e-292;
Matches 1630; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 14 GTCATTATTTAGAACGAGGTCTTGAATGACGAGATTAATCGGACCGGAGATACGTTG 73
Db 161 GTCATTATTTAGAACGAGGTCTTGAATGACGAGATTAATCGGACCGGAGATACGTTG 220
QY 74 AGCAATACAGGAATGACTACTGTGAGGATATGTTCTTAGACATTTATCACAGACATG 133
Db 221 AGCAATACAGGAATGACTACTGTGAGGATATGTTCTTAGACATTTATCACAGACATG 280
QY 134 AAAGCGGTATCGAATCCACTGCAGTAAATCTTCAGTCCGACGAGGAGAGACGTCCTA 193
Db 281 AAAGCGGTATCGAATCCACTGCAGTAAATCTTCAGTCCGACGAGGAGAGACGTCCTA 340
QY 194 AAAGGAGCGCAATAGACACATGTTTCAAGTCATCATGTCAGTCCGAGGAGCCACCGAAG 253
Db 341 AAAGGAGCGCAATAGACACATGTTTCAAGTCATCATGTCAGTCCGAGGAGCCACCGAAG 400
QY 254 AAAGTCCAGGATATGAGGATGATGAGGAGGTCACTCATCTGTCAAAGTGAGAGC 313
Db 401 AAAGTCCAGGATATGAGGATGATGAGGAGGTCACTCATCTGTCAAAGTGAGAGC 460
QY 314 TTCTTAAGAGCAAGATATGAATCGTGGACATTTTGGGTGAAGGACCTTTTGCAGAAAGTTG 373
Db 461 TTCTTAAGAGCAAGATATGAATCGTGGACATTTTGGGTGAAGGACCTTTTGCAGAAAGTTG 520
```


627	QY	TCAAATTTGACCACATCAGGAGATGGCGTATCAGATCTGCCAGTCAATAAATTTTTTACA	686
2449	Db	TCAAATTTGACCACATCAGGAGATGGCGTATCAGATCTGCCAGTCAATAAATTTTTTACA	2508
687	QY	TCATAATAATTAAACCCATACAGATCTGAAGCCTGAAATATTTTGTGTGAAGTCTGA	746
2509	Db	TCATAATAATTAAACCCATACAGATCTGAAGCCTGAAATATTTTGTGTGAAGTCTGA	2568
747	QY	CTATGTAGTCAAAATATAAATTTCTAAAAATGAAACGTGATGAACGCACACTGAAAAACACAGA	806
2569	Db	CTATGTAGTCAAAATATAAATTTCTAAAAATGAAACGTGATGAACGCACACTGAAAAACACAGA	2628
807	QY	TATCAAAGTTGTGACCTTTGGAAAGTGCACGATATGATGATGAACATCACAGTACTTTGGT	866
2629	Db	TATCAAAGTTGTGACCTTTGGAAAGTGCACGATATGATGATGAACATCACAGTACTTTGGT	2688
867	QY	GTCTACCCGCACTACAGAGCTCCGAGGTCATTTTGGCTTTAGTTGGTCTCAGCCTTG	926
2689	Db	GTCTACCCGCACTACAGAGCTCCGAGGTCATTTTGGCTTTAGTTGGTCTCAGCCTTG	2748
927	QY	TGATGTTTGGAGCATAGGTTGCATTTCTTATTTGAATATTTACCTTGGTTTTCACAGTCTTTCA	986
2749	Db	TGATGTTTGGAGCATAGGTTGCATTTCTTATTTGAATATTTACCTTGGTTTTCACAGTCTTTCA	2808
987	QY	GACTCATGATAGTAAAGACGACCTGGCAATGATGGAACGAAATATTAGGACCCCATACAACA	1046
2809	Db	GACTCATGATAGTAAAGACGACCTGGCAATGATGGAACGAAATATTAGGACCCCATACAACA	2868

Altechul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shcherchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalek, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 1914)

Strausberg, R.

Direct Submission

Submitted (15-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

Contact: amadan@systemsbiology.org

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 18 Row: C Column: 18

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6671765.

Location/Qualifiers

1. 1914

/organism="Mus musculus"

/mol_type="mRNA"

/strains="CZECH II"

/db_xref="taxon:10090"

/clone="MGC:13871 IMAGE:3982920"

/tissue_type="Mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMV."

/clone_lib="NCI CGAP_Lu29"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1. 1914

/gene="Clk4"

/codon_start=1

/product="CDC like kinase 4"

/protein_id="AAH12675.1"

/db_xref="GI:15215131"

/db_xref="LocusID:12750"

/db_xref="MGI:109851"

/translation="MRHSKRTCPDMSRESMGHESYSGSHKRRSHSSQTQNRHCK PHQFDSDCHLEARNLDRDRYIDIRNDYCEGYVFRHYRHDVESTYRHS KSVSRKSSPKRKNRCPASHQSHSKRRKRSIEDDEGLHICQSGDLRARYE IVDTLGAGFGKVECDHGMGLHVAVKIVKQVGRREARSEIQVLEHLNSTDPNS

ORIGIN

Query Match 56.6%; Score 1277.6; DB 10; Length 1914;
Best Local Similarity 87.5%; Pred. No. 6.5e-227;
Matches 1418; Conservative 0; Mismatches 200; Indels 3; Gaps 2;

QY 14 GTCAATTATTTAGAACAGAGTCCCTTGAATGACGAGATATTCGGACCGGAGATAGCTTG 73
DB |||||
QY 261 GTCACTATTTAGAACAGAGTCCCTTGAATGACGAGATATTCGGACCGGAGATACATG 320
DB |||||
QY 74 ACGAATACAGCAATGACTACTGTGAAGATATGTTCTTACGACATTTATCAGAGACATG 133
DB |||||
QY 321 ATGAATACAGCAATGACTACTGTGAAGATATGTTCTTACGACATTTATCAGAGACG 380
DB |||||
QY 134 AAAGCGGTATCGAATCCATCGCAGTAAATCTTCAGTCCGACGAGGAGAGCAGTCTCTA 193
DB |||||
QY 381 AAAGCACTTACCGGATCCATTCAGTAAATCTTCAGTCCGACGAGGAGAGCAGCCCTA 440
DB |||||
QY 194 AAAGGAGCGCAATAGACATCTGTTCAAGTCACTCAGTTCAGAGAGCCACCGAAGA 253
DB |||||
QY 441 AGAAGAGCGTAATAGACCTGTGCAAGTCACTCAGTCCGACGAGGAGGAGCCACCGA 500
DB |||||
QY 254 AAAGATCCAGGAGTATAGAGATGATGAGGAGGTCACCTGATCTGTCAAAGTGAGAGC 313
DB |||||
QY 501 AAAGATCCAGGAGTATAGAGATGATGAGGAGGTCACCTGATCTGTCAAAGTGAGAGC 560
DB |||||
QY 314 TTCTAAGAGCAAGATATGAAATCGTGGACACTTTGGGTGAAGAGCCCTTTGCGAAAGTTG 373
DB |||||
QY 561 TTCTAAGAGCAAGATATGAAATCGTGGACACTTTGGGTGAAGAGCCCTTTGCGAAAGTTG 620
DB |||||
QY 374 TAGAGTGCAATTCATGCGCATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGGCA 433
DB |||||
QY 621 TAGAGTGCAATTCATGCGCATGGCATGGCATGGCATGGCATGGCATGGCATGGCATGGCA 680
DB |||||
QY 434 TAGCGCGTTACCGTGAAGCAGCTCGTTCCAGAAATCCAAAGTATTAGAGCACTTAAATAGTA 493
DB |||||
QY 691 TAGCGCGTTACCGGAGGAGCAGCTCGTTCTGAATCCAAAGTATTGGAGCACTTGAACAGCA 740
DB |||||
QY 494 CTGATCCCAATAGTGTCTTCGAGTGTGTCCAGATGCTTGAAGTGGTTGATCATCTATGTC 553
DB |||||
QY 741 CTGACCCCAACAGTGTCTTCGAGTGTGTCCAGATGCTTGAAGTGGTTGATCATCTATGTC 800
DB |||||
QY 554 ATGTTTGTATTTGTTTGAACACTCTGGGACCTTAGTACTTACATTTTCAATTAAGAAAAA 613
DB |||||
QY 801 ATGTTTGTATTTGTTTGAAGCTGTGGGACCTTAGTACTTATTTTAAAGAAAAATA 860
DB |||||
QY 614 GCTTTCTGCCATTTCAAATTCACACATCAGGAGATGGCGTATCAGATCTGCCAGTCAA 673
DB |||||
QY 861 GTTTTCTGCCATTTCAAATTTGATCATCAGGCAATGGCTTATCAGATCTGCCAGTCTA 920
DB |||||
QY 674 TAAATTTTTCATCATATAATTAATTAACCCATACAGATCTCGAAGCTCGAAATATTTTGT 733
DB |||||
QY 921 TAAATTTTTCATCATATAATTAATTAACACACGAGCTTAAACCTGAAATATTTTAT 980
DB |||||
QY 734 TTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTTAAATTAAGAGCTGAACCCACAC 793
DB |||||
QY 981 TTGTGAAGTCTGACTATGTAGTCAAAATATAATTTCTTAAATTAAGAGCTGAACCCACAC 1040
DB |||||
QY 794 TGAATAACACAGATATCAAAAGTGTGTTGAGTGGAGTGCAAGCTGATCATGATGAACATC 853
DB |||||
QY 1041 TGAATAACACAGATATCAAAAGTGTGTTGAGTGGAGTGCAAGCTGATCATGATGAACATC 1100
DB |||||
QY 854 ACAGTACTTTTGTGCTTACCCCGGACATACAGAGTCCCGAGGCTCAATTTTGGCTTTAGGT 913
DB |||||
QY 1101 ATAGTACTTTTGTGCTTCAACAGGACATACAGGCTCCAGAGGCTCAATTTTGGCTTAGGT 1160
DB |||||
QY 914 GGTCTCAGCCTTGTGATGTTTGGAGCATAGGTTGATTTCTTATTTGAAATATTACCTTGGT 973
DB |||||

Db 1161 GGTCTCAGCCTTGTGATGTTTGGAGCATAGGCTGCAATCTTATTGAGTACTACCTTGGGT 1220
QY 974 TCACAGTCTTTTCAGACTCATGATAGTAAGAGCACCTGGCAATGATGAAGCAATATTAG 1033
Db 1221 TCACAGTCTTTTCAGACCCACGATAGTAAGAGCACCTGGCAATGATGAGCGGATCTTAG 1280
QY 1034 GACCCATACCAACAACATGATTTCAGAAAAACAAGAAAAACGCAAGTATTATTTCCACATAACC 1093
Db 1281 GACCCATCCAGCACATATGATCCAGAGAACAAGAAAAACGCAAGTATTATTTCCACATAACC 1340
QY 1094 AGCTAGATGGAGTAAACACAGTCTTCTGCTGATGATATGTTAGAGAGACGCTGCAACCGT 1153
Db 1341 AGCTAGATGGAGCAGCATAGTTCAGCTGGAGATATGTTAGAGAGCGCTGCAACCGT 1400
QY 1154 TGAAGGATTTATGCTTTGTCATGATGAAGACATGAGAACTGTTTGCACCTGTTCCAA 1213
Db 1401 TAAAGGAATTTATGCTGTCATGACGAAGAGCATGAGAACTGTTTGCACCTGTTCCAA 1460
QY 1214 GAATGTTAGAAATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCAATTCAGCATCCTT 1273
Db 1461 GATGTTGGAGTATGACCCAGGAGAGGATCACCTTGGATGAAGCAATTCAGCATCCTT 1520
QY 1274 TCTTGACTTATTAAGAAAGAAATGAAATGGAATCAGTGTCTTATATATATCTCT 1333
Db 1521 TCTTGACTTATTAAGAAAGAAATG-AGTGGAGTCAAGGCTCTTCTGTGACTTCTCTA 1579
QY 1334 AGAGAGATTAATTAAGACTGTGTCAGTCAACTAAGCAATTTCTAATATTTTGTAAACATT 1393
Db 1580 GGAGCAGTACTTCCAGACTGTGTCAGTCAACTAAGCAATTTCTAATATTTTGTAAACATT 1639
QY 1394 AAATATTTTGTACAGTTAAAGTGAATATTTGATGTTTGTATCAATAGCAATTAATTAAC 1453
Db 1640 AAATATTTTGTACAGTTAAATGTTAAATCTGATGTTTGTATCAATAGCAATTAATTAAC 1699
QY 1454 TTGTTAAGCAAGTATGCTTGAATGAATGATAGAAAAATTAATTAATTTTCTTTT- 1512
Db 1700 TTGTTAAGTATGCTTGAATGAATGATAGAAAAATTAATTAATTTTCTTTT- 1759
QY 1513 -TGAATTTACCAATTTTAAATACCTTTGAATATCTTTGTCAGTGAATATGAT 1571
Db 1760 ATGTTAATGTCGCTTTAAAGCTTTTGAATGCTTTTGTGTCAGTGAATATGAT 1819
QY 1572 TGATCTGCTCTTTGTACATGAGGTGACCTCTGAAGTGAATTTTGTGAGTAAAGGAA 1631
Db 1820 TGGTCCCATCTTTGTACATGAAGTGTGACTCTGAAGTGAATTTTGTGCTCAAGTAA 1879
QY 1632 A 1632
Db 1880 A 1880

RESULT 13
AX961894 1881 bp DNA linear PAT 14-JAN-2004
LOCUS Sequence 105 from Patent WO03104277.
DEFINITION AX961894
ACCESSION AX961894
VERSION AX961894.1 GI:40881304
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE Sugahara,T., Matsuda,A., Honda,G., Muramatsu,S. and Ishizawa,K.
AUTHORS Stat6 activation gene
TITLE Patent: WO 03104277-A 105 18-DEC-2003;
JOURNAL Asahi Kasei Kabushiki Kaisha (JP)
FEATURES Location/Qualifiers
1..1881
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
90..1535
CDS

/note="unnamed protein product"
/codon_start=1
/protein_id="CAP06456.1"
/db_xref="GI:40881305"
/translation="MRSHKTHCPDWDRESWGHSYSGSHKRRKSHSHSTQENRHK
PHQKDSCHYLEARCLNDRYIDYENDYCEGVPRHYHRRDVESTYRHK
KSSVRSRSPKRRKRRPCASHQSHSRKRSRSTEDDEEHLICOGDVLRRYE
IVDTLGGAGTKVVEICIDHGDGLHVAIKVKNVYREARSEIQVLEHNSDPNS
VFCVOMLEFDDHGHVICIVFELLGLSTYDFIKENSLFQIDHIRMVAVGTCOSINF
LHNKLPHDLKPEHLLFKSDYVVKYNSKMRDERTLKNTDIKVDFGSAATYDEHH
STLSTRHYRAPEVILALHNSQPCDWSIGCILLEYLIGTIVFOTHSKHELMNERI
LGP1PAHMIQKTRKRYPHNQLDDEHSSAGRYVRRRKLKELFELCHDEHKLFD
LVRRMLEYDPARRITLDELQHPFFDLKRX"

ORIGIN
Query Match 56.6%; Score 1277.4; DB 6; Length 1881;
Best Local Similarity 87.9%; Pred. No. 7.1e-227; Indels 3; Gaps 2;
Matches 1413; Conservative 0; Mismatches 192;
QY 14 GTCATTATTAGAAGCAAGTCTTGAATGAGCGAGATTATCGGACCGGAGATACGTTG 73
Db 250 GTCATTATTAGAAGCAAGTCTTGAATGAGCGAGATTATCGGACCGGAGATACATTG 309
QY 74 ACGAATACAGGAATGACTTGTGAAGGATATGTTCTTAGACATTATCAGAGACATTG 133
Db 310 ATGAATACAGAAATGACTTGTGCAAGGATATGTTTCCAAGACATTACCATAGAGACGTTG 369
QY 134 AAAGCGGTATCGAATCCACTGCGAGTAACTTCCAGTCCGAGCAGGAGAGAGCGATCCTA 193
Db 370 AAAGCATTTACCGGATCAATTCGAGTAACTTCCAGTCCGAGCAGGAGAGAGAGCGATCCTA 429
QY 194 AAAGAGAGCGAATAGACACTGTTCAAGTCAATCAGTCACTGTTCAAGAGCAACCGAAGGA 253
Db 430 AGAGAGAGCGTAAATAGACCTTGTGCAAGTCACTGTTCAAGAGCAACCGAAGGA 489
QY 254 AAAGATCCAGAGTATAGAGGATGATGAGAGGGTCACTGATCTGTCAAAGTGGAGACG 313
Db 490 AAAGATCCAGAGTATAGAGGATGATGAGAGGGTCACTGATCTGTCAAAGTGGAGACG 549
QY 314 TTCTAAGAGCAAGATATGAAATCGTGGACACTTTTGGGTGAAGGAGCGCTTTGCGCAAGTTG 373
Db 550 TTCTAAGAGCAAGATATGAAATCGTGGACACTTTTGGGTGAAGGAGCGCTTTGCGCAAGTTG 609
QY 374 TAGAGTCAATTGATCATGGCATGAGTGCATGTCATGTAGCAGTGAATAATCGTAAATAATG 433
Db 610 TAGAGTCAATTGATCATGGCATGAGTGCCTTACATGTAGCAGTGAATAATCGTAAATAATG 669
QY 434 TAGGCGGTTACCGTGAAGCAGCTGTTCAAGAAATCCAAAGTATTAGAGCACTTAATAGTA 493
Db 670 TAGGCGGTTACCGGAGGAGCAGCTGTTCTGAAATCCAAAGTATTGAGAGCACTTGAAACGCA 729
QY 494 CTGATCCCAATAGTGTCTTCCGATGTGTCAGATGCTAGAAATGTTTGCATCATGCTGC 553
Db 730 CTGATCCCAATAGTGTCTTCCGATGTGTCAGATGCTAGAAATGTTTGCATCATGCTGC 789
QY 554 ATGTTTGTATTGTTGTTGAACCTACTCTGGGACTTAGTACTTACGATTTCAATTAAGAAAAACA 613
Db 790 ATGTTTGTATTGTTGTTGAGCTGCTGGGACTTAGTACTTACGATTTCAATTAAGAAAAACA 849
QY 614 GCTTTCTGCCATTTCAAAATTCACACATCAGGAGATGCGGTATCAGATCTGCCAGTCAA 673
Db 850 GTTTTCTGCCATTTCAAAATTCACATCAGGAGAAATGGCTTATCAGATCTGCCAGTCTA 909
QY 674 TAAATTTTTTACATCATATAATAATTAACCCATACAGATCTGAAGCTCGAAATATTTTGT 733
Db 910 TAAATTTTTTACATCATATAATAATTAACACACGAGCTTAACCTGAGAAATATTTTAT 969
QY 734 TTGTGAAGTCTGACTATGATGTAATAATAATTTCAAAATGAAGCTGATGAACCCACAC 793
Db 970 TTGTGAAGTCTGACTATGATGTAATAATAATTTCAAAATGAAGCTGATGAACCCACAC 1029
QY 794 TGAAGAAACAGATATCAAGTGTGTTGAGCTTTGGAGTGCACGATGATGATGAACATC 853

Db 1030 TGAATAACACAGATATCAAGTCTTGATTTTGGAGTGCAACATATGACGACGAAATC 1089
QY 854 ACAGTACTTTGTGTCTACCCGCACTACAGAGCTCCGAGGTCAATTTTGGCTTTAGGTT 913
Db 1090 ATAGTACTTTGTGTCTCAACAGGCACTACAGGCTCCAGAGGTCAATTTTGGCTCTAGGTT 1149
QY 914 GGTCTCAGCTTTGTGATGTTTGGAGCATAGGTTGATCTTATTGAAATTAACCTTGGTT 973
Db 1150 GGTCTCAGCTTTGTGATGTTTGGAGCATAGGTTGATCTTATTGAGTACTACCTTGGTT 1209
QY 974 TCACAGTCTTTTCAGACTCATGATAGTAAAGAGCACTGCGCAATGATGAACCAATATTAG 1033
Db 1210 TCACAGTCTTTTCAGACTCATGATAGTAAAGAGCACTGCGCAATGATGAACCAATATTAG 1269
QY 1034 GACCCATACCAACACATGATTTTCAGAAAACGCAAGTATTTTTCACCAATAACC 1093
Db 1270 GACCCATACCAACACATGATTTTCAGAAAACGCAAGTATTTTTCACCAATAACC 1329
QY 1094 AGCTAGATGGATGAACACAGTCTTCTGCTGTGATGATGTTAGGAGAGCTGCAACCGT 1153
Db 1330 AGCTAGATGGAGCAGCATAGTTCAGCTGGGAGATATGTTAGGAGAGCTGCAACCGT 1389
QY 1154 TGAAGGAATTTATGCTTTGTCTATGATGAAGACATGAGAACTGTTGACCTGGTTCGAA 1213
Db 1390 TAAAGGAATTTATGCTTTGTCTATGATGAAGACATGAGAACTGTTGACCTGGTTCGAA 1449
QY 1214 GAATGTTAGATATGATCCAACTCAAGAAATTTACCTTGGATGAAGCATTTGACGATCCTT 1273
Db 1450 GAATGTTGAGTATGACCCAGCAGAGGATCACCTTGGATGAAGCATTTGACGATCCTT 1509
QY 1274 TCTTTGACTTATTAAGAAATGAATGGAAATGAGTATGATGTTGATGATGATGATGATGAT 1333
Db 1510 TCTTTGACTTATTAAGAAATGAATGGAAATGAGTATGATGTTGATGATGATGATGATGAT 1568
QY 1334 AGAAGAGATTACTTAAGACTGTGCTAGTCACTAATCACTTCTAATATTTTGTAAACATT 1393
Db 1569 GGAGCAGTTACTTTCAGACTGTGCTAGTCACTAATCACTTCTAATATTTTGTAAACATT 1628
QY 1394 AAATTTATTTGATCACTTAAGTAAATTAATTTGATGTTTGTATCAATAGCATATTAAC 1453
Db 1629 AAATTTATTTGATCACTTAAGTAAATTAATTTGATGTTTGTATCAATAGCATATTAAC 1688
QY 1454 TTGTTAAGCAAGTATGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1512
Db 1689 TTGTTAAGTATGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1748
QY 1513 TGAATTTACCAATTTTAAATACCTTTGAAATATCTTTGCTCCAGTATTAATGATGAT 1571
Db 1749 ATGTTAAATGTCACCTTTTAAAGCTTTTAAATGCTTTTGTGTCAGTATTAATGATGAT 1808
QY 1572 TGATCTTGCTTTTGTACATGAGGTGACCTCTGAGAGTATTTTGTGTTT 1619
Db 1809 TGGTCCCATCTTTGTACATGAGGTGACCTCTGAGAGTATTTTGTGTTT 1856

RESULT 14
BD260616
LOCUS 1222 bp DNA linear PAT 17-JUL-2003
DEFINITION 49 human secreted proteins.
ACCESSION BD260616
VERSION BD260616.1 GI:33070386
KEYWORDS JP 2002539776-A/16.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1222)
AUTHORS Rosen,C.A., Ruben,S.M. and Komatsoulis,G.
TITLE 49 human secreted proteins
JOURNAL Patent: JP 2002539776-A 16 26-NOV-2002;
HUMAN GENOME SCIENCES INC
OS Homo sapiens (human)
PN JP 2002539776-A/16

PD 26-NOV-2002
PF 16-MAR-2000 JP 2000606614
PR 23-MAR-1999 US 60/126051,10-DEC-1999 US 60/169906 PI
CRAIG A ROSEN, STEVEN M RUBEN, GEORGE KOMATSOUIS PC
C12N15/09, A61K35/36, A61K38/00, A61K48/00, A61P17/02, A61P25/14, PC
A61P25/16,
PC A61P25/18, A61P25/24, A61P25/28, A61P29/00, C07K14/47, C07K16/18,
PC C12N1/15,
PC C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12Q1/02, C12Q1/68, G01N33/
53,
PC G01N33/566 G01N33/68, C12N1/5/00, A61K37/02, C12N5/00 CC 49
human secreted proteins
FH Key Location/Qualifiers
FT source 1.1222
FT source /organism="Homo sapiens (human)".
FEATURES
source
1.1222
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 52.5%; Score 1185; DB 6; Length 1222;
Best Local Similarity 99.9%; Pred. No. 1e-209;
Matches 1196; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 446 GTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGAGCACTTAAATAGTACTGATCCCAATA 505
Db 9 GTGAAGCAGCTCGTTTCAGAAATCCAAAGTATTAGAGCACTTAAATAGTACTGATCCCAATA 68
QY 506 GTGCTTCCGATGTGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 565
Db 69 GTGCTTCCGATGTGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 128
QY 566 TGTGTTGAATCTGCGACTTACTTACGATTTTCAATTAAGAAACAGCTTCTGCGCAT 625
Db 129 TGTGTTGAATCTGCGACTTACTTACGATTTTCAATTAAGAAACAGCTTCTGCGCAT 188
QY 626 TTCAATTTGACACATCAGGAGATGCGGTATCAGATCTGCGAGTCAATAAATTTTAC 685
Db 189 TTCAATTTGACACATCAGGAGATGCGGTATCAGATCTGCGAGTCAATAAATTTTAC 248
QY 686 ATCAATAAATAAACCACATCAGATCTGAAAGCTGAAAAATATTTGTTGTAAGTCTG 745
Db 249 ATCAATAAATAAACCACATCAGATCTGAAAGCTGAAAAATATTTGTTGTAAGTCTG 308
QY 746 ACTATGATGATCAATATTAATTTCTAAATGAAACGTGAAACGACACCTGAAAAACACAG 805
Db 309 ACTATGATGATCAATATTAATTTCTAAATGAAACGTGAAACGACACCTGAAAAACACAG 368
QY 806 ATATCAAAAGTTGTTGACTTTGCAAGTCAACGATGATGATGATGATGATGATGATGATGAT 865
Db 369 ATATCAAAAGTTGTTGACTTTGCAAGTCAACGATGATGATGATGATGATGATGATGATGAT 428
QY 866 TGTCTACCCGCACTACAGAGCTCCGAGGTCAATTTGCGCTTTAGGTTGGTCTCAGCCTT 925
Db 429 TGTCTACCCGCACTACAGAGCTCCGAGGTCAATTTGCGCTTTAGGTTGGTCTCAGCCTT 488
QY 926 GTGATGTTTGAGCATAGGTTGCAATCTTATTAATTAATTAATTAATTAATTAATTAATTAAT 985
Db 489 GTGATGTTTGAGCATAGGTTGCAATCTTATTAATTAATTAATTAATTAATTAATTAATTAAT 548
QY 986 AGACTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1045
Db 549 AGACTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 608
QY 1046 AACACATGATTTTCAGAAAAACAAGAAACGCAAGTATTTTTCACCAATACCGCTAGATTGGG 1105
Db 609 AACACATGATTTTCAGAAAAACAAGAAACGCAAGTATTTTTCACCAATACCGCTAGATTGGG 668
QY 1106 ATGAACACAGTTCTGCTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1165
Db 669 ATGAACACAGTTCTGCTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 728

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2005, 23:59:12 ; Search time 41 Seconds
(without alignments)
1044.304 Million cell updates/sec

Title: US-10-801-671-2

Perfect score: 2410

Sequence: 1 MCIPLEASHSVEEDTHPSHY.....QRITLDEALQHFFDLKKK 445

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1887	78.3	484	2 S33641	protein kinase clk
2	1837	76.2	483	2 A39676	protein kinase STY
3	1727	71.7	454	2 A38643	protein kinase EC
4	1417	58.8	499	2 S33637	protein kinase clk
5	1321	54.8	490	2 S33639	protein kinase clk
6	1317	54.6	490	2 S70352	protein kinase Dar
7	1222	50.7	517	2 A54099	protein kinase Dar
8	965.5	40.1	903	2 T19209	probable protein k
9	876	36.3	306	2 I49068	protein kinase STY
10	848.5	35.2	427	2 T05560	protein kinase AFC
11	839	34.8	431	2 T04125	protein kinase PKI
12	830	34.4	467	2 S71169	protein kinase, 54
13	790	32.8	400	2 T04460	protein kinase AME
14	767	31.8	575	2 JC7794	lammer kinase homo
15	767	31.8	690	2 T38052	probable protein k
16	718.5	23.8	737	1 S64767	probable serine/th
17	530	22.0	543	2 S62456	probable serine-th
18	524	21.7	1087	2 S58147	protein kinase - f
19	523	21.7	508	2 T22440	hypothetical prote
20	523	21.7	817	2 T22442	hypothetical prote
21	504.5	20.9	589	2 JG0195	protein kinase DYE
22	478.5	19.9	629	2 JG0195	protein kinase DYE
23	477	19.8	1157	2 C96761	hypothetical prote
24	474	19.7	948	2 T24445	hypothetical prote
25	472.5	19.6	754	2 JC4898	Down-syndrom-crit
26	448.5	18.6	1189	2 T17088	homeodomain-intera
27	443.5	18.4	1155	2 B96761	probable protein k
28	441	18.3	1192	2 T17089	homeodomain-intera
29	437	18.1	642	2 T45904	protein kinase-lik

probable protein k
serine/threonine p
protein kinase YAK
hypothetical prote
protein kinase YAK
homeodomain-intera
mRNA splicing-asso
protein B0464.5a l
hypothetical prote
hypothetical prote
probable protein k
protein kinase hom
protein B0464.5c l
protein kinase (SC
serine protein kin
mitogen-activated

ALIGNMENTS

RESULT 1
S53641
protein kinase clk1 (EC 2.7.1.1) - human
C:Species: Homo sapiens (man)
C>Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C:Accession: S53641
R:Hanes, J.; von der Kammer, H.; Klaudiny, J.; Scheit, K.H.
J. Mol. Biol. 244, 665-672, 1994
A>Title: Characterization by cDNA cloning of two new human protein kinases. Evidence by
A:Reference number: S53637; MUID:95082033; PMID:7990150
A:Accession: S53641
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-484 <HAN>
A:CROSS-references: UNIPROT:P49759; GB:L29219; NID:G632963; PIDN:AA61480.1; PID:G632963
C:Superfamily: human protein kinase clk1; protein kinase homology
C:Keywords: alternative splicing; phosphotransferase
F:159-430/Domain: protein kinase homology <KIN>

Query Match	78.3%	Score	1887	DB 2	Length	484			
Best Local Similarity	82.1%	Pred. No.	2.1e-83						
Matches	352	Conservative	30	Mismatches	45	Indels	2	Gaps	2
QY	18	SHYLEARSINERDRYRDEYNDYCEGVPRHYRDIHSGYRIHCKSSVRRSSP	77						
DB	54	SHYLESRINEKDYHSRYIDEYNDYQGCPEGRQDRHSRYQNHSSKSSGRSSY	113						
QY	78	KRK-RNRHCSSH-QSRKSHRRKRSRSTEDDEGHLCQSGDVLARARVEIVDTLGEAGF	135						
DB	114	KSKRIHHSTSHRRSHGSKSHRRKTRSVEDDEGHLCQSGDVLARSARYEIVDTLGEAGF	173						
QY	136	KVCEIDHGMGMHVAVKIKVNGVRYEAAARSEIQVLEHLNSTDPNSVFCVQMLEWFDH	195						
DB	174	KVCEIDHKGAGRVAVKIKVNDRYCEAAARSEIQVLEHLNTDPNSTFRVQMLEWFEH	233						
QY	196	HGHVCIYFELLGLSTYDFIKENSFLPQIDHROIWAYOIQSINFLHKNLTHDTLKPEN	255						
DB	234	HGHVCIYFELLGLSTYDFIKENGLPFLDHRKWAYOIQSINFLHKNLTHDTLKPEN	293						
QY	256	ILFKVSDYVVKYNSKMKRDERTLNKTDIKVDFGSATYDDHHSTLVSTRYRAPEVILA	315						
DB	294	ILFKVSDYVVKYNSKMKRDERTLNKTDIKVDFGSATYDDHHSTLVSTRYRAPEVILA	353						
QY	316	LGMQPCDVMWIGICILLYYLGFTVFQTHDSKEHLAMMERILGPIPOHMIQKTRKRYFH	375						
DB	354	LGMQPCDVMWIGICILLYYLGFTVFQTHDSKEHLAMMERILGPIPOHMIQKTRKRYFH	413						
QY	376	HNQLDWDEHSSAGRYVRRRKPLKEFMKLDVRRMLEYDPTQRTITLDEALQ	435						
DB	414	HNQLDWDEHSSAGRYVRRRKPLKEFMKLDVRRMLEYDPTQRTITLDEALQ	473						
QY	436	HPFFDLLKK	444						

C;Date: 21-Feb-1992 #sequence_revision 21-Feb-1992 #text_change 09-Jul-2004

C;Accession: A38643

C;Johnson, K.W.; Smith, K.A.

J. Biol. Chem. 266, 3402-3407, 1991

A;Title: Molecular cloning of a novel human cdc2/CDC28-like protein kinase.

A;Reference number: A38643; MUID:91139618; PMID:1704889

A;Accession: A38643

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-454 <JOH>

A;Cross-references: UNIPROT:Q9NRL6; GB:M59287

C;Genetics:

A;Gene: GDB:CLK

A;Cross-references: GDB:126862

A;Map position: 7q31-7q31

C;Superfamily: human protein kinase

C;Keywords: phosphotransferase

F;129-400/Domain: protein kinase homology <KIN>

Query Match 71.7%; Score 1727; DB 2; Length 454;

Best Local Similarity 76.6%; Pred. No. 8.5e-76;

Matches 327; Conservative 26; Mismatches 46; Indels 28; Gaps 3;

Qy 18 SHYLEARSLNERDYDRRYVDEYRNDYCEGVPRHYHRDIESGVRIHCSKSSVRSRSP 77

Db 54 SHYLESRSINEKYHSRRYIDEYRNDYTCGCEPGHRRDHSRYQNHSSKSSGSGRSSY 113

Qy 78 KRKNRHCSSHQSRSKSHRRKRSRSIEDDEBEGHLICQSGDVLRLARYEIVDTLGEAGFKV 137

Db 114 KSKRIH-----HSTSHRSH-----GD-----EIVDTLGEAGFKV 145

Qy 138 VECIDHGMGNHVAKIVKNGVRYREARAEISIQVLEHLNSTDPNSVRCQMLEWFDHGG 197

Db 146 VECIDHKGAGRHVAKIVKNGVRYCEAARSEIQVLEHLNTDPNSTFRCQMLEWFEHHG 205

Qy 198 HVCIVFELLGLSTYDFIKENSLFPQIDHIROMAYQICQSFNHLHNKLTHTDLKPENIL 257

Db 206 HICIVFELLGLSTYDFIKENGLFPLDHIKRMAYQICKSVNHLNKLTHTLKPENIL 265

Qy 258 FVKSIDYVVKNSKMKRDETLKNTDIKVPDGSATYDDEHSTLSTVSTRHYRAPEVILALG 317

Db 266 FVQSDYTEAYNPVKIKRDETLNPDIKVDPGSATYDDEHSTLSTVSTRHYRAPEVILALG 325

Qy 318 WSPQCDVWSIGICILIEYLGFTVQTHDSKEHLAMMERILGPIQHMIOKTRKRYFHN 377

Db 326 WSPQCDVWSIGICILIEYLGFTVPTTHDSKEHLAMMERILGPKHMLQTRKRYFHD 385

Qy 378 OLDWDEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLVRRLMLEYDPTQRTITLDEALQHP 437

Db 386 RLDWDEHSSAGRYVSRACKPLKEFMLSQDVEHERLFDLIQKMLEYDPAKRITLREALKHP 445

Qy 438 FFDLLKK 444

Db 446 FFDLLKK 452

RESULT 4

S53637

protein kinase clk2, long splice form (EC 2.7.1.1) - human

C;Species: Homo sapiens (man)

C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004

C;Accession: S53637; T08025

R;Hanes, J.; von der Kammer, H.; Klaudiny, J.; Scheit, K.H.

J. Mol. Biol. 244, 665-672, 1994

A;Title: Characterization by cDNA cloning of two new human protein kinases. Evidence by

A;Reference number: S53637; MUID:95082033; PMID:7990150

A;Accession: S53637

A;Molecule type: mRNA

A;Residues: 1-499 <HAN>

A;Cross-references: UNIPROT:P49760; GB:I29218; NID:G632967; PIDN:AAA61482.1; PID:G632968

R;Winfield, S.L.; Tayebi, N.; Martin, B.M.; Gimms, E.I.; Sidransky, E.

Genome Res. 7, 1020-1026, 1997

C;Species: Homo sapiens (man)

A;Title: Identification of three additional genes contiguous to the glucocerebrosidase 1c

Db 474 HPFFDLLKK 482

RESULT 2

A39676

protein kinase STY (EC 2.7.1.1) - mouse

N;Alternate names: protein kinase clk

C;Species: Mus musculus (house mouse)

C;Date: 08-Nov-1991 #sequence_revision 08-Nov-1991 #text_change 09-Jul-2004

C;Accession: A39676; S13364

R;Howell, B.W.; Afar, D.E.H.; Lew, J.; Douville, E.M.J.; Icely, P.L.E.; Gray, D.A.; Bell

Mol. Cell. Biol. 11, 568-572, 1991

A;Title: STY, a tyrosine-phosphorylating enzyme with sequence homology to serine/threonine

A;Reference number: A39676; MUID:91094875; PMID:1986248

A;Accession: A39676

A;Molecule type: mRNA

A;Residues: 1-483 <HOW>

A;Cross-references: UNIPROT:P22518; GB:M38381; NID:G201070; PIDN:AAA40151.1; PID:G201071

R;Ben-David, Y.; Letwin, K.; Tannock, L.; Bernstein, A.; Pawson, T.

EMBO J. 10, 317-325, 1991

A;Title: A mammalian protein kinase with potential for serine/threonine and tyrosine pho

A;Reference number: S13364; MUID:91122038; PMID:1825055

A;Accession: S13364

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-378, P', 380-483 <BEN>

A;Note: the sequence from Fig. 2 is inconsistent with that shown in Fig. 1 in having 448

C;Superfamily: human protein kinase clk1; protein kinase homology

C;Keywords: phosphotransferase; serine/threonine-specific protein kinase

F;158-429/Domain: protein kinase homology <KIN>

Query Match 76.2%; Score 1837; DB 2; Length 483;

Best Local Similarity 79.7%; Pred. No. 5.1e-81;

Matches 349; Conservative 32; Mismatches 51; Indels 6; Gaps 4;

Qy 9 HSVEETHPSHYLEARSLNERDYDRRYVDEYRNDYCEGVPRHYHRDIESGYRIHCSKS 68

Db 48 HSKTTD--SYLESRSINEKAYHSRRYVDEYRNDY-GYEPGHPYGPGRSGRYQMHSSKS 103

Qy 69 SVRSRRSPRK-RNK-HCSHQSRKSHRRKRSRSIEDDEBEGHLICQSGDVLRLARYEIV 126

Db 104 SGRSGRSSYKSHRRHHTSQHSHGKSHRRKRSRSVEDDEBEGHLICQSGDVLRLARYEIV 163

Qy 127 DTIGEGAFKGVCEIDHGMGNHVAKIVKNGVRYREARAEISIQVLEHLNSTDPNSVRC 186

Db 164 DTIGEGAFKGVCEIDHGMGNHVAKIVKNGVRYCEAARSEIQVLEHLNTDPNSTFRC 223

Qy 187 VQMLEWFDHGHVCIIVFELLGLSTYDFIKENSLFPQIDHIROMAYQICQSFNHLHNKLT 246

Db 224 VQMLEWFEHRGHICIVFELLGLSTYDFIKENSLFPQIDHIROMAYQICKSVNHLNKL 283

Qy 247 THTLKPENILFVKSIDYVVKNSKMKRDETLKNTDIKVPDGSATYDDEHSTLSTVSTRH 306

Db 284 THTLKPENILFVKSIDYTEAYNPVKIKRDETLNPDIKVDPGSATYDDEHSTLSTVSTRH 343

Qy 307 YRAPEVILALGWSQCDVWSIGICILIEYLGFTVQTHDSKEHLAMMERILGPIQHMIOQ 366

Db 344 YRAPEVILALGWSQCDVWSIGICILIEYLGFTVFTSTHDSREHLAMMERILGPKHMIQ 403

Qy 367 KTRKRYFHNLQDWDDEHSSAGRYVRRCKPLKEFMLCHDEHEKLFDLVRRLMLEYDPTQ 426

Db 404 KTRKRYFHNLQDWDDEHSSAGRYVRRCKPLKEFMLSQDAEHELLFDLIQKMLEYDPAK 463

Qy 427 RITLDEALQHPFDLLKK 444

Db 464 RITLKEALKHPFFYPLKK 481

RESULT 3

A38643

protein kinase (EC 2.7.1.37) cdc2/cdc28-like - human

C;Species: Homo sapiens (man)

A:Reference number: Z16482; MUID:97474796; PMID:9331372

A:Accession: T08825

A:Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-499 <WIN>

C:Cross-references: GB:AF023268; NID:G2564910; PIDN:AA651817.1; PID:G2564911

C:Comment: The short splice form of this protein (see PIR:S53638) lacks the protein kinase

C:Genetics:

A:Gene: clk2

C:Function:

A:Description: Phosphorylates Ser, Thr, and Tyr residues on proteins in the spliceosomal

C:Superfamily: human protein kinase clk1; protein kinase homology

C:Keywords: alternative splicing; phosphotransferase

F:161-440/Domain: protein kinase homology <KIN>

Query Match 58.8%; Score 1417; DB 2; Length 499;

Best Local Similarity 59.7%; Pred. No. 5.5e-61;

Matches 259; Conservative 69; Mismatches 98; Indels 8; Gaps 3;

18 SHYLEARS-----LNERDYRRRYVDEY-RNDYCEGYVPRHYRDIESGYRHCSKSVR 71

50 SYHVRSSSYDRSDRRYDRYCGSYRRNDYSDRDGAYDYTDYRHSYEQRENSYR 109

72 SRRSPK--RKRNHCSHQSRKSHRRKRSRSIEDDEGHLICQSGDVLARVEIVDTL 129

110 SQRSRRKRRRRRSRTFSRSSQHSRRRAKSVEDDAEGHLIYHVGDLQERYEIVSTL 169

130 GEGAFGKVECTDHGMGHWAVKIVKNGVRYEAARSEIOVLEHLNSTDPNSVRCVQM 189

170 GEGTFRVQCVDHRRGGARVAKIKVVEKKEARLEINVLKINEKDPDKMLCQVM 229

190 LEWFDHGHVCITVFELGLSTYDFIKENSFLPFQIDHIROMAYQICQSFNHLHNKLTHT 249

230 FDFPDYHGMCTSFELGLSTDFDLKNNLYLPYTHQVRHMAFOLCQAVKFLHDNKLHT 289

250 DLKPNILFVKSIVVYKNSKMRDERTIKNTDVKVDFGATYDDEHSTLVSTRHYRA 309

290 DLKPNILFVNSDYELTYNLEKRDERSVKSTAVRVDFGATFDEHSTLVSTRHYRA 349

310 PEVILALGWSQPCDWSIGCIIIEYLGFTVQTHDSKEHLAMMERIIGPIPQHMIOKTR 369

350 PEVILELGWSQPCDWSIGCIIIEYVGTFTLQTHDNREHLAMMERIIGPIPSMIRKTR 409

370 KRKYFHHNLQDWEHSSAGRYVRRCKPKFELMCHDBEHEKLFDLVRRMLEYDPTQRT 429

410 KQKYFGRGLDWDENTSAGRYVRENCKPLRLYITSEAEHQLFDLIESMLEYEPKELT 469

430 LDEALQHPFFDLK 443

470 LGEALQHPFFARLR 483

RESULT 5

S53639

protein kinase clk3 (EC 2.7.1.-) - human

C:Species: Homo sapiens (man)

C>Date: 28-Oct-1995 #sequence_revision 07-Feb-1997 #text_change 09-Jul-2004

C:Accession: S53639; S71040

R:Hanes, J.; von der Kammer, H.; Klaudiny, J.; Scheit, K.H.

J. Mol. Biol. 244, 665-672, 1994

A:Title: Characterization by cDNA cloning of two new human protein kinases. Evidence by

submitted to the EMBL Data Library, January 1995

A:Reference number: S53637; MUID:95082033; PMID:7990150

A:Accession: S53639

A:Molecule type: mRNA

A:Residues: 1-490 <HAN>

C:Cross-references: UNIPROT:P49761; GB:L29217

R:Hanes, J.J.; der Kammer, H.; Klaudiny, J.J.; Scheit, K.H.

submitted to the EMBL Data Library, January 1995

A:Reference number: S71040

A:Accession: S71040

A:Molecule type: mRNA

A:Residues: 1-131, 'TG', 134-490 <HAW>

A:Cross-references: EMBL:L29217; NID:G632971; PIDN:AAA61484.1; PID:G632972

C:Superfamily: human protein kinase clk1; protein kinase homology

C:Keywords: alternative splicing; phosphotransferase

F:154-433/Domain: protein kinase homology <KIN>

Query Match 54.8%; Score 1321; DB 2; Length 490;

Best Local Similarity 55.3%; Pred. No. 2e-56;

Matches 251; Conservative 60; Mismatches 101; Indels 42; Gaps 6;

QY 4 PLEASHSVEEDTHPSHYLEARSLENERDYRRRYVDEYRNDYCEGYVP-----R 51

Db 43 PPRSRSRSHDRLP-----YQRRYRRRSDSYR---CEERSPSFGEDYDGSSRSR 90

QY 52 HYHRDIESG-YRI-----HCSKSVRRSRSPKRNHCHSHQSRKSHRRKRKRSIED 105

Db 91 HRRSRERGYPYTRKHAHCHK-----RRTSCSSASSRSQQSSKRSRSVED 138

QY 106 DEEGHLICQSGDVLARVEIVDTLGEAGFVGKVEICIDHGMGHWAVKIVKNGVRYEAA 165

Db 139 DKEGHLVCRIGDWQERYEIVGNLGEFTGKVEICIDHARKSQVALKIIRNVGKTRAA 198

QY 166 RSEIQVLEHLNSTDPNSVRCVQMLEWFDHGHVCIVFELLGLSTYDFIKENSFLPFQID 225

Db 199 RLEINVLLKIKEKKNKFLCVLMSDFWENFGHMCIAFELLGKNTFEFLKKNFPQYPLP 258

QY 226 HIROMAYOICQSFNHLHNKLTHTDLKPNILFVKSIVVYKNSKMRDERTIKNTDVK 285

Db 259 HVRHMAVQLCHALRFLHENQLTHTDLKPNILFVNSEPETLYNEHKSCEKSVKNTSIRV 318

QY 286 VDFGSATYDDEHSTLVSTRHYRAPEVILALGWSQPCDWSIGCIIIEYLGFTVFO 345

Db 319 ADFGSATFDEHHTTIVATRHYPPEVILELGMWQPCDWSIGCILFYEGFTLFTO 378

QY 346 SKEHLAMMERIIGPIPQHMIOKTRKRYFHHNQLDWEHSSAGRYVRRCKPKLEFMLCH 405

Db 379 NREHLVMEKILGPIPSMIRKTRKQKYFYKGLVWDENSDGRVYKENCPLKSYMLQD 438

QY 406 DEEHEKLFDLVRRMLEYDPTQRTITLDEALQHPFF 439

Db 439 SLEHVQLFDLRRMLEFDPQAQRITLAEALLHPFF 472

RESULT 6

S70352

protein kinase clk3 (EC 2.7.1.-), testis-specific - rat

N:Alternate names: LAMMER kinase

C:Species: Rattus norvegicus (Norway rat)

C>Date: 21-Apr-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004

C:Accession: S70352

R:Becker, W.; Kentrup, H.; Heukelbach, J.; Joost, H.G.

Biochim. Biophys. Acta 1312, 63-67, 1996

A:Title: cDNA cloning and characterization of rat Clk3, a LAMMER kinase predominately ex

A:Reference number: S70352; MUID:96271481; PMID:8679717

A:Accession: S70352

A:Molecule type: mRNA

A:Residues: 1-490 <BEC>

C:Cross-references: UNIPROT:Q63117; EMBL:X94351; NID:g1149536; PIDN:CAA64076.1; PID:g1149

C:Genetics:

A:Gene: clk3

C:Superfamily: human protein kinase clk1; protein kinase homology

C:Keywords: phosphotransferase; protein kinase

F:154-433/Domain: protein kinase homology <KIN>

Query Match 54.6%; Score 1317; DB 2; Length 490;

Best Local Similarity 55.1%; Pred. No. 3.2e-56;

Matches 250; Conservative 61; Mismatches 101; Indels 42; Gaps 6;

QY 4 PLEASHSVEEDTHPSHYLEARSLENERDYRRRYVDEYRNDYCEGYVP-----R 51

Db 43 PPRSRSRSHDRIP-----YQRRYRRRSDSYR---CEERSPSFGEDYDGSSRSR 90

QY 52 HYHRDIESG-YRI-----HCSKSVRRSRSPKRNHCHSHQSRKSHRRKRKRSIED 105

Db 91 HRRSRERGYPYTRKHAHCHK-----RRTSCSSASSRSQQSSKRSRSVED 138

224	QY	IDHIROMAYOICOSINFLPHNKU TH DLK PENIL FVKS DYV--VKY-NSKWRD---BRT	277
197	Db	IDL VRE IGW QLLE CVAF MDH LM TH DLK PEN ILLV SS DDY VKI PEY KGS RLQ DD VCY KRV	256
278	QY	LKNTD IKVND FGSAT YDD DEH HS TV STR HYR APE VIL ALG WSQ CDVMS IGC ILE YVLG	337
257	Db	PKS SA KV IDFG STT YR QD TV STR HYR APE VIL GLG WSY PCDVMS VGC ILV ELCTG	316
338	QY	FTV FQ THDS KEH LAMMER ILG PIQ PHMI QKT--RKRKY F HNN QLD DWD SHS AGR VRRRC	395
317	Db	EAL FQ THEN LEH LAMMER VLG PP PQ QML KKV DRHSE KYVVRGR LDW PDGAT SRD SLKAVL	376
396	QY	K--PLKEFM LCH--DEB EKL FDL VRRM LEYD PQ RT L DEALQ HPFF	439
377	Db	KL PR QNL LMOH VDH SAGEL INN VOGL RF DPS ERIT AREAL RPFF	421

RESULT 11
T04125
protein kinase PK12 (EC 2.7.1.-), ethylene-induced - common tobacco
C/Species: *Nicotiana tabacum* (common tobacco)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 16-Aug-2004
C/Accession: T04125
R/Sesaa, G.; Raz, V.; Savaldi, S.; Fluhr, R.
Plant Cell 8, 2223-2234, 1996
A/Title: PK12, a plant dual-specificity protein kinase of the LAMMER family, is
A/Reference number: Z15228; MUID:97143872; PMID:8989879
A/Accession: T04125
A/Status: translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 1-431 <SES>
A/Cross-references: UNIPROT:O49967; EMBL:U73937; NID:G2911279; PIDN:AAC04324.1
A/Experimental source: cultivar Samsun NN
C/Genetics:
A/Gene: PK12
C/Superfamily: protein kinase homology
C/Keywords: ATP; phosphotransferase; protein kinase
F:94-381/Domain: protein kinase homology <KIN>

Query Match	34.8%;	Score	839;	DB	2;	Length	431;
Best Local Similarity	39.9%;	Pred. No.	1.7e-33;				
Matches	175;	Conservative	80;	Mismatches	130;	Indels	54;
Gaps	10						
QY	29	RDYDRRYDEYNDYCEGVVPHYHDISSGVRHCSKSSVRSR-----	73				
DB	11	RPYMDRRPKRPPLDDPSHTPK-AOSGIYVGQEVGNSSSYVHSRLLPDHDLSLVYKGLAQ	69				
QY	74	RSSPKKRNHRCSSHQSRKSRSHKRSRSTEDDEEGHLICQSGDVLNARYEIVDTLGECA	133				
DB	70	KSGPPRR-----DDDKGHYMFELGNLTTRYKILKKIGSGT	106				
QY	134	FGKVRECDIHGMDGMHVAVKIVKNVGRYREAAARSEIQVLEHLNSTDPNSVRCFQOMLEWF	193				
DB	107	FGQVLECDWDEQKGF-VAIKIIRSIKKYREAAVVDVLQLLGRYDRGGT-RCVQLRNWF	164				
QY	194	DHGHGTCVIFELLGLSTYDFIKENSFLFPFOIDHIRQWAYQICQISFLHNKLTHTDLKP	253				
DB	165	DYRNHITCLVFEKLGSLDFDLRKNRYAFVVDLVREITGRQLELCAVFMHDMRLIHTHTDLKP	224				
QY	254	ENILFYKSDVY----VKYNSKMKRD-----ERTIKNTDIKVVDFGSATVYDDEHSHSLVSTR	305				

225	ENLFVSADYIIKVPDVKGTW	SHRDRSFSKRLPKSSAIK	VIDFGSTAYERP	DHNYIVSTR	284
306	HYRAPEVILALGWSQPCD	VMSIGCIIETYYLGT	VFQTHDSKEHLAMMER	ILGPIQHRM	365
285	HYRAPEVILGLGWSYPCD	LMSVGCIIETELCS	GEALFQTHENLEHLAMMER	IVGLPSPQML	344
366	OKTRK--RXVFHENO	LDDWDSHSSAGRV	RRRCK--PLKEFWLCH-	DEEHEKFLDLVRML	420
345	KYVDRAEAKYVRGR	LDDWPEGATISRESIK	SVKMLPRLONVQHV	DHDSAGDLIDLLOGL	404
421	KYVDPTQRIITLDEAL	QHPFF	439		

Db 405 RFPDSIRMTAHDALRHPFF 423

RESULT 12

S71169

protein kinase, 54K (EC 2.7.1.1) - Arabidopsis thaliana

N:Alternate names: protein F4P12.270

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004

C:Accession: S71169; T45897

R:Kurumori, T.; Yamamoto, M.

submitted to the EMBL Data Library, January 1995

A:Description: A.thaliana genes encoding protein kinases of a new family.

A:Reference number: S71169

A:Accession: S71169

A:Molecule type: mRNA

A:Residues: 1-467 <KUR>

A:Cross-references: UNIPROT:P51566; EMBL:D45354; NID:G6421132

R:Bloeker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, P.; Salanoubat, M.

submitted to the Protein Sequence Database, January 2000

A:Reference number: Z23016

A:Accession: T45897

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-467 <BLO>

A:Cross-references: EMBL:AL132966

A:Experimental source: cultivar Columbia; BAC clone F4P12

C:Genetics:

A:Map position: 3

A:Introns: 56/3; 115/1; 124/3; 175/2; 191/3; 230/1; 260/3; 313/1; 334/3; 367/2; 402/3

A:Note: F4P12.270

C:Superfamily: human protein kinase ckl1; protein kinase homology

C:Keywords: ATP; phosphotransferase; protein kinase

F:113-401/Domain: protein kinase homology <KIN>

Query Match 34.4%; Score 830; DB 2; Length 467;

Best Local Similarity 45.8%; Pred. No. 5e-33;

Matches 160; Conservative 71; Mismatches 102; Indels 16; Gaps 8;

QY 105 DDEGHLICQSGDVLRYEIVDTLGEAGFGKVVCEIDHGMGMHVAIVKNGVRYREA 164

DB 97 DDKGHYVFFVVDLTPRYQILSKMGEGTFGVLECDWDRDTK-EYVAIKIIRSIKRYDAA 155

QY 165 ARSEIQVLEHLNSTDPNSVFCVQMLEWFDHGHGVCIVFELLGLSTYDFIKENSFLPFQI 224

DB 156 AMEIDVLQLTRHDVGG-SRCVQIRNWFEDYRNHICIVEKLGPSLYDFLRKNSYRSFPI 214

QY 225 DHIQWVQICOSINFLHNKLTHTDLPENILFVKSDYV----VKYNSKMKRDERTLKN 280

DB 215 DLVRELGLLESVAVMHDLRLHTDLPENILVSSSEYIKIPDYKFLSRPTKDGSYFKN 274

QY 281 ----TDIKVDFGSGATYDEHSTLSTVSTHYRAPEVILALGWSQPCDVWSIGCILLIEYYL 336

DB 275 LPKSAIKLIDFGSTTFEQQDNYIVSTHYRAPEVILGVGNVPCDLWSIGCILLIELCS 334

QY 337 GFTVQTHDSKEHLAMERILGPIQHMIOKT--RRKYFHHN-QLDWDHSSAGRYVR 393

DB 335 GEAIFQTHENLEHLMERVLGFLPPHVLADRSEKVFRRGAKLDWPEGATSDSLKA 394

QY 394 --RCRKLKEFMILCH-DEEHEKLFDLVRRMLEYDPTQRTILDEALQHPFF 439

DB 395 VVKLPRLPNLMQHVHDSAGDLIDLLQLGLRYDPTFERFKAREALNHPFF 443

RESULT 13

T04460

protein kinase AME3 (EC 2.7.1.1) - Arabidopsis thaliana

N:Alternate names: protein F4P11.140

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 16-Aug-2004

C:Accession: T04460

R:Bevan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Ansoorge, W.; Hoheisel, J.; Mewes, H.W.

submitted to the Protein Sequence Database, April 1998

A:Reference number: Z15360

A:Accession: T04460

A:Molecule type: DNA

A:Residues: 1-400 <BEV>

A:Cross-references: UNIPROT:P51568; EMBL:AL022537

A:Experimental source: cultivar Columbia; BAC clone F4D11

C:Genetics:

A:Gene: AME3

A:Map position: 4

A:Introns: 30/3; 71/1; 80/1; 131/2; 147/3; 186/1; 216/3; 266/1; 287/3; 320/2; 355/3

A:Note: F4D11.140

C:Superfamily: protein kinase homology

C:Keywords: ATP; phosphotransferase; protein kinase

Query Match 32.8%; Score 790; DB 2; Length 400;

Best Local Similarity 41.7%; Pred. No. 3.5e-31;

Matches 169; Conservative 70; Mismatches 134; Indels 32; Gaps 11;

QY 57 IESGYRIHCSKSSVRSR-----RSSPKRKR---NRHCSHOSRSKSHRKRGRSRSIED 105

DB 2 IANGFE-SMDKERVKRPRMTWDEAPAEPAKRAVIKGGSDGRILSPPLR-----DD 53

QY 106 DEEGHLICQSGDVLRYEIVDTLGEAGFGKVVCEIDHGMGMHVAIVKNGVRYREA 165

DB 54 DRDGHYVFLSDNLTTPRYKILSKMGEGTFGVLECDWDRDTK-EYVAIKIIRSIKRYDAA 112

QY 166 RSEIQVLEHLNSTDPNSVFCVQMLEWFDHGHGVCIVFELLGLSTYDFIKENSFLPFQID 225

DB 113 MIEDVLQKLVSKDKGRT-RCVQKQWFDYRNHICIVEKLGPSLYDFLRKNSYSAFPLA 171

QY 226 HIRQWVQICOSINFLHNKLTHTDLPENILFVKSDYVVKYNSKMKRDERT-----LKN 280

DB 172 LVRDFGQCLLESVAVMHDLQVHTDLPENILVSSSENVKLPDNKRSAAETHFRCLPKS 231

QY 281 TDIKVDFGSGATYDEHSTLSTVSTHYRAPEVILALGWSQPCDVWSIGCILLIEYYLGFV 340

DB 232 SAIKLIIDFGSTVCDNRHHSIVOTRHYRSPSEVILGLWSYQCDLWSIGCILLIEYLGAL 291

QY 341 FQTHDSKEHLAMERILGPIQHMIOKT--RRKYFHHN-QLDWDHSSAGRYVR--RRC 395

DB 292 FQTHDNLHLEHLMERVLGFLPEHMTKASRGAEKFFRGCRINWPEGANSRESIRAVKEL 351

QY 396 KPLKEFMILCH-DEEHEKLFDLVRRMLEYDPTQRTILDEALQHPFF 439

DB 352 DRLKDMVSKVDNTRSRPADLLYGLLAYDPSERLTANEALDHPFF 396

RESULT 14

JC7794

lammer kinase homolog protein 1, Lkh1 protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 02-Apr-2002 #sequence_revision 02-Apr-2002 #text_change 09-Jul-2004

C:Accession: JC7794

R:Kim, K.H.; Cho, Y.M.; Kang, W.H.; Kim, J.H.; Byun, K.H.; Park, Y.D.; Bae, K.S.; Park, I.

Biotech. Biophys. Res. Commun. 289, 1237-1242, 2001

A:Title: Negative regulation of filamentous growth and flocculation by Lkh1, a fission yeast, by repressing expression of cell surface proteins for flocculation and morphogenesis

A:Reference number: JC7794

A:Accession: JC7794

A:Molecule type: mRNA

A:Residues: 1-575 <KIM>

A:Cross-references: UNIPROT:Q10156; GB:AF334941

A:Experimental source: strain SP286, ED665

C:Comment: This protein, a homolog of the dual-specificity protein kinase of the LAMMER family, by repressing expression of cell surface proteins for flocculation and morphogenesis

C:Genetics:

A:Gene: Lkh1+

Query Match 31.8%; Score 767; DB 2; Length 575;

Best Local Similarity 41.5%; Pred. No. 6.1e-30;

Matches 141; Conservative 78; Mismatches 115; Indels 6; Gaps 3;

QY 105 DDEGHLICQSGDVLRYEIVDTLGEAGFGKVVCEIDHGMGMHVAIVKNGVRYREA 164

Db	229	DDDDGHYKVVVPNSKANRYTVVRLGHCTGCKVLCQVDQS--TGRHCAIKVTRAIPIKYREA	281
Qy	165	ARSEQTQVLEHLNSTDPNSVFRVCQWLEWFDHHGHVCIVFELLGLGSTVYDFPKENSFIPFOI	224
Db	288	SHIEURLVQTIASHSDPTNENKCIQLRDYDFYRKHCIVTDLFGWSVDFDLKNNYIPFPL	347
Qy	225	DHIQWMAVYQIQOSINFLHNKLTHTDLKPENILFKVSDYVVKNSKMKRDERLTAKNTDK	284
Db	348	KHIQMLSQQLFKSVAFLHSLGLVHTDLKPENVLLVSNASRTIRLPYRNSQKVLSNCEIR	407
Qy	285	VDFGSGATVDDHHSTLVSTRHYRAPEVILALGWSQPCDVMSTGCIILIEYLGFTVPQTH	344
Db	408	LIDFGSATPDEYSHSVSVSTRHYRAPEITILGLGWSYPCDVMSTGCIILVELFTQQAIFQTH	467
Qy	345	DSKEHLAMMERITGLPIPOHMI---QKTRKRYFHHNOLDWDEHSSAGRYVR--RRCKPLK	399
Db	468	EUSEHLCKMEKILGFDFDRNMLRSRSRTSQRFKSDGKRYPLNLTWPKSNINYLSQLOTUE	527
Qy	400	EFWMLCHDEBEKLFDLVRMLLEYDPTQBITLDEALQHPFF	439
Db	528	QIFAVSSPEVALLDLKKVYFDPKRRITAKEALWHPFF	567

```

RESULT 15
T38052
probable protein kinase - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
C/Accession: T38052
R/Lye: G.; Churchell, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, February 1995
A/Reference number: Z21765
A/Accession: T38052
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-690 <LYE>
A/Cross-references: EMBL:Z69239; PIDN:CAA93220.1; GSPDB:GN00066; SPDB:SPAC1D4.11c
A/Experimental source: strain 972h-; cosmid c1d4
C/Genetics:
A/Gene: SPDB:SPAC1D4.11c
A/Map position: 1
C/Superfamily: yeast probable protein kinase KNS1; protein kinase homology

```

Query Match	31.8%;	Score 767;	DB 2;	Length 690;
Best Local Similarity	41.5%;	Pred. No. 7.1e-30;		
Matches	141;	Conservative 78;	Mismatches 115;	Indels 6;
Gaps	3;			
Qy	105	DDEGHLICQSGDVLRA	YEIVDTLGGAGFGKVV	ECIDHGMDGWHVAVKIVKNCVRYEA 164
Db	344	DDDGHYKVPNKFANRY	TVRLLGHGTFGKICQY	DQS--TGRHCAIKVTRAIPIKRYEA 402
Qy	165	ARSEIQVLEHLNSTD	PNSVFCVQWLEWDFD	HGHGHCIVFELGLGLSTYDFIKENSPFLPQI 224
Db	403	SLIELRVLIQIAHSD	PTNENKCIQLRDYFDY	RKHICIVTDLFGSVDFPLKNNVYIPL 462
Qy	225	DHIQWAIQCQSINFL	HNHKNLTHTDLPENIL	FKVSDYVVKYNSKMRDERTLKNTDIK 284
Db	463	KHIQMLSQQLFKSVA	FLHSLGLVTDLPENVL	LVSNASRTIRLPYRNSQVLNCSCEIR 522
Qy	285	VWDFGSATYDDEHST	LVSTRHYRAPEVILAL	GWNSQPCDVWSIGCILLEYLGVFTVFOTH 344
Db	523	LIDFGSATFDEYHSS	VVSTRHYRAPEIILG	WNSYPCDVWSIGCILFELFTGQALFQTH 582
Qy	345	DSXEHLLAMMERILG	PIPOHMI--QKTRKRY	FTHNQLDWDHDSAGRYVR--RRCKPLK 399
Db	583	EDSEHLCWMEKILGF	PDENMLRSRSTSORF	FKSDGKVRYPLSNTPKSINVLSQLTLE 642
Qy	400	EFMLCHDEHEKLPDL	VRMLLEYDPTQRIITL	DEALQHPFF 439
Db	643	QIFAVSSPEVALLDL	LKKVFPYDPPKRRIT	AKAELWHPFF 682

Search completed: March 13, 2005, 00:08:47

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2005, 00:05:12 ; Search time 29 Seconds
(without alignments)
726.868 Million cell updates/sec

Title: US-10-801-671-2

Perfect score: 2410

Sequence: 1 MCIPLEASHSVEDTHPSHY.....QRITLDEALQHFFDLKKK 445

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 190900 seqs, 47368951 residues

Total number of hits satisfying chosen parameters: 190900

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents_AA_New.*
1: /cgn2_6/ptodata/1/paa/pct_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US11_NEW_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query		Length	DB	ID	Description
	Score	Match				
1	1692	70.2	374	1	PCT-US04-30360-133	Sequence 133, App
2	1314.5	54.5	490	6	US-10-450-763-53713	Sequence 53713, A
3	1310	54.4	490	6	US-10-878-630A-2	Sequence 2, Appli
4	473.5	19.6	421	1	PCT-US04-30360-134	Sequence 134, App
5	472.5	19.6	763	1	PCT-US04-42360-2024	Sequence 2024, Ap
6	451.5	18.7	429	1	PCT-US04-30360-135	Sequence 135, App
7	420	17.4	1043	7	US-11-050-926-258	Sequence 258, App
8	419.5	17.4	371	1	PCT-US04-30360-136	Sequence 136, App
9	380	15.8	373	1	PCT-US04-30360-35	Sequence 35, Appl
10	337	14.0	703	6	US-10-450-763-46352	Sequence 46352, A
11	335.5	13.9	1030	1	PCT-US05-03526-8	Sequence 8, Appli
12	326.5	13.5	358	1	PCT-US04-30360-60	Sequence 60, Appl
13	326.5	13.5	359	7	US-11-021-951-187	Sequence 187, App
14	326.5	13.5	480	6	US-10-450-763-44291	Sequence 44291, A
15	323.5	13.4	359	1	PCT-US04-30360-17	Sequence 17, Appl
16	322.5	13.4	360	1	PCT-US04-42360-1761	Sequence 1761, Ap
17	322.5	13.4	360	1	PCT-US04-42360-2389	Sequence 2389, Ap
18	322	13.4	355	1	PCT-US04-30360-19	Sequence 19, Appl
19	322	13.4	358	1	PCT-US04-42360-2165	Sequence 2165, Ap
20	321	13.3	360	1	PCT-US04-30360-56	Sequence 56, Appl
21	321	13.3	360	1	PCT-US04-42360-2167	Sequence 2167, Ap
22	319.5	13.3	326	1	PCT-US04-30360-15	Sequence 15, Appl
23	319.5	13.3	326	1	PCT-US04-30360-52	Sequence 52, Appl
24	316	13.1	304	7	US-11-033-515-13	Sequence 13, Appl
25	316	13.1	304	7	US-11-033-515-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1

PCT-US04-30360-133

; Sequence 133, Application PC/TUS0430360

; GENERAL INFORMATION:

; APPLICANT: PLEXIKON, INC.

; TITLE OF INVENTION: MOLECULAR SCAFFOLDS FOR KINASE LIGAND DEVELOPMENT

; FILE REFERENCE: 039363-1703

; CURRENT APPLICATION NUMBER: PCT/US04/30360

; CURRENT FILING DATE: 2004-09-15

; PRIOR APPLICATION NUMBER: 60/503,277

; PRIOR FILING DATE: 2003-09-15

; NUMBER OF SEQ ID NOS: 167

; SOFTWARE: PatentIn Ver. 3.2

; SEQ ID NO 133

; LENGTH: 374

; TYPE: PRT

; ORGANISM: Unknown Organism

; FEATURE:

; OTHER INFORMATION: Description of Unknown Organism: Mammalian

; OTHER INFORMATION: protein sequence

PCT-US04-30360-133

Query Match	70.2%	Score 1692;	DB 1;	Length 374;
Best Local Similarity	84.7%	Pred. No. 1.4e-101;		
Matches 315;	Conservative 24;	Mismatches 31;	Indels 2;	Gaps 2;
QY	75	SSPKRK-RNRHCSH-QSRKSHRRKRSRSTEDDEGHLCQSGDVLKRYEIVDTLGG 132		
Db	1	SSYSKRIHHSTSHRRSHGSHKRRKTRSVDEDEGHLCQSGDVLKRYEIVDTLGG 60		
QY	133	AFGKVECDHGMGMHVAVKIVKNGRYREARSEIQVLEHLSNTDPSNFRVCQMLEW 192		
Db	61	AFGKVECDHKGGRHVAVKIVKNDRYCEARSEIQVLEHLSNTDPSNFRVCQMLEW 120		
QY	193	PDHGHVCIIVPELLGLSTYDIKENSFLPFQIDHIRQWAYQICOSINFLHKNLTHTLK 252		
Db	121	FERHGHICIVPELLGLSTYDIKENGFLPFPLDHIRKWAYQICKSVNFLHKNLTHTLK 180		
QY	253	PENILFVKSDVVKVYKNSKMRDERTLKNQTDIKVDFGSAATYDDHSHSLVSTRHYRAPEV 312		
Db	181	PENILFVQSDTTEAYNPKIKRDETLNPDILKVDVDFGSAATYDDHSHSLVSTRHYRAPEV 240		
QY	313	ILALGWQPCDWSIGCIIIEYLGFTVFQTHDSKEHLAMMERILGPIPOHMIQKTRKK 372		
Db	241	ILALGWQPCDWSIGCIIIEYLGFTVFQTHDSKEHLAMMERILGPIPKMIQKTRKK 300		
QY	373	YFHNQLDWDHSHSAGRYVRRCKPLKPFMLCHDEEHKFLDLVRMLSEYDPTQRTILDE 432		

Db 301 YPHDRDLWDHDSAGRYVSRACKFLKBFMLSQDVEHERLFDLIQKMLEYDPAKRIITLRE 360

QY 433 ALQHFFFDLLKK 444

Db 361 ALKHFFFDLLKK 372

RESULT 2

US-10-450-763-53713

; Sequence 53713, Application US/10450763

; GENERAL INFORMATION:

; APPLICANT: Hyseg, Inc

; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

; FILE REFERENCE: 790CIP3/US

; CURRENT APPLICATION NUMBER: US/10/450,763

; CURRENT FILING DATE: 2003-06-11

; PRIOR APPLICATION NUMBER: PCT/US01/08631

; PRIOR FILING DATE: 2001-03-30

; PRIOR APPLICATION NUMBER: 09/540,217

; PRIOR FILING DATE: 2000-03-31

; PRIOR APPLICATION NUMBER: 09/649,167

; PRIOR FILING DATE: 2000-08-23

; NUMBER OF SEQ ID NOS: 60736

; SOFTWARE: Custom

; SEQ ID NO 53713

; LENGTH: 490

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: DOMAIN

; LOCATION: (271)..(290)

; OTHER INFORMATION: TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE domain identified

; OTHER INFORMATION: by eMATRIX, accession number PR00109B, p-value=7.559e-10, raw score

; OTHER INFORMATION: of 12.27

; FEATURE:

; NAME/KEY: DOMAIN

; LOCATION: (155)..(470)

; OTHER INFORMATION: Eukaryotic protein kinase domain identified by Pfam,

; OTHER INFORMATION: accession name pkinase, E-value=4.8e-69, Pfam score of 242.8

US-10-450-763-53713

Query Match

Best Local Similarity 54.5%; Score 1314.5; DB 6; Length 490;

Matches 251; Conservative 59; Mismatches 89; Indels 21; Gaps 7;

QY 34 RRYVDE---YRNDYCEGVVPR-HYHRDIESGRIHCSKSSVSRSSPKRKR-----NR 83

Db 64 RYHSSEGRSGSYCEHRSRKHQRSS-----WSSSSDRTTR---RRRDSYHVR 114

QY 84 HCSSHQSRKS-HRRKRSRSDDEDEHGLICQSGDVLRYEIVDTLGEAGFGKYVECID 142

Db 115 RCSRFSRSSQHSRKAQSVEDDTEGHLIYHVGDLQERYEIVSTLGKGTGRVVQCVD 174

QY 143 HGMDGMHVAIKVNGYRRAARSEIOVLEHLNSTDPNSVRCVOMLEWFDHGHVCIV 202

Db 175 HRRRGARVALKIKNVEKYKEAARLEIKVLEKINEKDPGKQL-CVQMFDFDYGHCMS 233

QY 203 FELLGLSTYDFIKENSLPFOIDHIRMAYOICQSFNLHKNLTHTDLKPNILFVKSD 262

Db 234 LELGLSTFDLKNHNLPIYHQVHHWASQCAVQKFLHDKLHTDLKPNILFVNSD 293

QY 263 YVVKYNSKWRDERTLKNLTKVDVFGSATYDDEHSTLVSTRHYRAPEVILALGWSQPC 322

Db 294 YELTYNLEKKHRSVKSTAVRVGDFGSAFDHEHHSITVSTRHYRAPEVILELGSQPC 353

QY 323 DVWSIGCILIIYYLGLTFVQTHDSKEHLAMMERILGPIDPMIOKTRKRYFHHNQDWD 382

Db 354 DVWSIGCIIFYYVGLTFQTHDRNQHATMERILGPIPSRMIRKTRQKYFYRGLDWD 413

QY 383 EHSAGYVRRRCRCKPLKEFMLCHDEHEKLFDLVRMLEYDPTORTITLDEALQHPFFDL 442

Db 414 ENTISAGYVRENCKPLROYLTISEAEDHQLFDLIESMLEYPAQRLTUGEALQHPFFSRL 473

RESULT 3

US-10-878-630A-2

; Sequence 2, Application US/10878630A

; GENERAL INFORMATION:

; APPLICANT: Wisotzkey, Robert G.

; TITLE OF INVENTION: CDC28-Like protein kinase CLK3 disruptions, compositions and

; FILE REFERENCE: R-3217/40338.0241USU1

; CURRENT APPLICATION NUMBER: US/10/878,630A

; CURRENT FILING DATE: 2004-06-28

; PRIOR APPLICATION NUMBER: US 60/483,017

; PRIOR FILING DATE: 2003-06-26

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 2

; TYPE: PRT

; LENGTH: 490

; ORGANISM: Mus musculus

US-10-878-630A-2

Query Match 54.4%; Score 1310; DB 6; Length 490;

Best Local Similarity 55.6%; Pred. No. 5.4e-77;

Matches 248; Conservative 62; Mismatches 110; Indels 26; Gaps 4;

QY 4 PLEASHSVEEDTHPSHYLEARSLNERDYDRRYVDEYRNDYCEGVVPRHYHRDIESGYRI 63

Db 43 PPRSRSRSHDRIP-----YQRRVREHRDSDTYR---CEERSPSFG---EDCYGS 86

QY 64 HCSKSVSRSSPKR-----KRNHCSSHQSRKSHRRKRSSIEDDEGHLIC 113

Db 87 SRSRHRRSRERAPYRTRKHAHHCHKRTRSCSSASSRSQSSKRSRSDDEKGEHLVC 146

QY 114 OSGDVLRYEIVDTLGEAGFGKYVECIDHGMDGMHVAIKVNGYRRAARSEIOVLE 173

Db 147 RIGDWLQERYEIVGNLGEFTGFKVVECLDHARGSKQVALKIRNNGKTRRARLEINLVK 206

QY 174 HLNSTDPSVRCVOMLEWFDHGHVCIVFELLGLSTYDFIKENSLPFOIDHIRMAYQ 233

Db 207 KIKEKXENKFLCVLMSDFNFGHMCIAFELLGKNTFEFLKENNFQYPLPHVRHMAVQ 266

QY 234 ICQSFNLHKNLTHTDLKPNILFVKSDYVVKYNSKWRDERTLKNLTKVDVFGSATY 293

Db 267 LCHALFLHENQLTHTDLKPNILFVNSDFETLYNEHKSCKEKSVKNTSIRVADFGSATF 326

QY 294 DDEHSTLVSTRHYRAPEVILALGWSQPCDVWSIGCILIIYYLGLTFVQTHDSKEHLAM 353

Db 327 DHEHTTIVATRYRPPVEVILELGAQPCDVWSIGCILFEYIRGFTLFTQTHENREHLVMM 386

QY 354 ERILGPIPOHMIQTRKRYFHHNQDWDHSSAGRYVRRRCRCKPLKEFMLCHDEHEKLF 413

Db 387 EKILGPIPSHMIHRTRKQKYFGKGLVMDENSSDGRVYVKENCKPLKSYMLQDSLEHVQLF 446

QY 414 DLVRMLEYDPTORTITLDEALQHPFF 439

Db 447 DLMRMLEFDPAQRITLAEALLHPFF 472

RESULT 4

PCT-US04-30360-134

; Sequence 134, Application PC/TUS0430360

; GENERAL INFORMATION:

; APPLICANT: PLEXIKON, INC.

; TITLE OF INVENTION: MOLECULAR SCAFFOLDS FOR KINASE LIGAND DEVELOPMENT

; FILE REFERENCE: 039363-1703

; CURRENT APPLICATION NUMBER: PCT/US04/30360

; CURRENT FILING DATE: 2004-09-15

; PRIOR APPLICATION NUMBER: 60/503,277

; PRIOR FILING DATE: 2003-09-15

; NUMBER OF SEQ ID NOS: 167

; SOFTWARE: PatentIn Ver. 3.2

; SEQ ID NO 134

; LENGTH: 421

QY 428 ITLDEALQHPF 438
Db 368 ITPIETLNHPF 378

RESULT 7
US-11-050-926-258
; Sequence 258, Application US/11050926
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHWANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATZ
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/11/050,926
; CURRENT FILING DATE: 2005-02-04
; PRIOR APPLICATION NUMBER: US/10/097,340
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 258
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-050-926-258

Query Match 17.4%; Score 420; DB 7; Length 1043;
Best Local Similarity 26.2%; Pred. No. 1.2e-19;
Matches 126; Conservative 91; Mismatches 183; Indels 80; Gaps 16;

QY 1 MCIPLEASHSVEEDTH-----PSHYLEARSINERDYYRRVYDEVRNDYCEGYVPRVHR 55
Db 597 MSVPSEPS-SPQSSRTSPSPDDILERVAADVKEY-ERENVDTF----- 639
QY 56 DIESGYRTHCSKSSVRSSRSPKRR-----NRHCSSSHOSRSKSHRR--KRS 100
Db 640 --EASVKAHNLMTVEQNGSSQKLLAPDMFTSDDMFAAYFDSARLARAGIGKDFKEN 697
QY 101 RSIED---DEGHILICQSGDVLRYEIVDTLGEAGFGKVVCEIDHGMGMHVAVKI 157
Db 698 PNLNRDNTDAEGYRVNIGEVLDKRYNYGYTGQGVFSNVVRARDNARANQEVAVKI 757
QY 158 VGRYREARSEIQVLEHLNSTDPNSVFCVQMLEWFDHGHVCI VFELLGLSTYDFIKE- 216
Db 758 NEMQKTGLKELEFLKLNADADPDDKFHCLRLFRHFYHKQHLCLVFEPLSNLREVLKXY 817

QY 217 NSFLPFQIDHIRMAYQICQSFNLFHNNKLTHTLKPENILFVKSDDYVVKNSKMKRDER 276
Db 818 GKDVGLHIKAVRSYSQQLFLALKKRCNHLHADIKPNILNVNES----- 862
QY 277 TLKNTDIKVDVFGSATY--DDEHSTLVSTRHYRAPEVILALGWSQPCDVMSIGCILLEY 334
Db 863 ---KTILKLCDFGSASHVADNDITPYLVS-RFYRAPEIIGKSYDYDGIDMWSVGCTLYEL 918
QY 335 YLGTFTVQTHDSKEHLAMMERILGPIQHMIOK-TRREKYFHHN-----QLDWEDEHSSAGR 389
Db 919 YTGKILFPKTNHNLKLANDLKGKMPNKMIRKGVFKDQDFDQNLNFMVIEVDKVTREK 978
QY 390 Y-VRRRCCKPLKEFM-----LCHDEHE--KLFDLVRRLMEYDPTORTILDEALQHPF 438
Db 979 VTVMSTNPTKDLADLIGCORLPEDQKVKVQLKDLLDQILMLDPKAKRISINQALQHPF 1038

RESULT 8
PCT-US04-30360-136
; Sequence 136, Application PC/TUS0430360
; GENERAL INFORMATION:
; APPLICANT: PLEXIKON, INC.
; TITLE OF INVENTION: MOLECULAR SCAFFOLDS FOR KINASE LIGAND DEVELOPMENT
; FILE REFERENCE: 039363-1703
; CURRENT APPLICATION NUMBER: PCT/US04/30360
; CURRENT FILING DATE: 2004-09-15
; PRIOR APPLICATION NUMBER: 60/503,277
; PRIOR FILING DATE: 2003-09-15
; NUMBER OF SEQ ID NOS: 167
; SOFTWARE: Patent In Ver. 3.2
; SEQ ID NO 136
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Mammalian
; OTHER INFORMATION: protein sequence
PCT-US04-30360-136

Query Match 17.4%; Score 419.5; DB 1; Length 371;
Best Local Similarity 29.5%; Pred. No. 4.6e-20;
Matches 107; Conservative 73; Mismatches 142; Indels 41; Gaps 10;

QY 98 KRSRSIED---DEGHILICQSGDVLRYEIVDTLGEAGFGKVVCEIDHGMGMHVAVKI 154
Db 23 KENPNLRDNTDAEGYRVNIGEVLDKRYNYGYTGQGVFSNVVRARDNARANQEVAVKI 82
QY 155 VKNVGRYREARSEIQVLEHLNSTDPNSVFCVQMLEWFDHGHVCI VFELLGLSTYDFI 214
Db 83 IRNNELMQKTGLKELEFLKLNADADPDDKFHCLRLFRHFYHKQHLCLVFEPLSNLREVL 142
QY 215 KE-NSFLPFQIDHIRMAYQICQSFNLFHNNKLTHTLKPENILFVKSDDYVVKNSKMKR 273
Db 143 KKYGVGLHIKAVRSYSQQLFLALKKRCNHLHADIKPNILNVNES----- 190
QY 274 DERTLKNTDIKVDVFGSATY--DDEHSTLVSTRHYRAPEVILALGWSQPCDVMSIGCIL 331
Db 191 -----KTILKLCDFGSASHVADNDITPYLVS-RFYRAPEIIGKSYDYDGIDMWSVGCTL 243
QY 332 IEYLGFTVQTHDSKEHLAMMERILGPIQHMIOK-TRREKYFHHN-----QLDWEDEHSS 386
Db 244 YELGTGKILFPKTNHNLKLANDLKGKMPNKMIRKGVFKDQDFDQNLNFMVIEVDKYTE 303
QY 387 AGRY-VRRRCCKPLKEFM-----LCHDEHE--KLFDLVRRLMEYDPTORTILDEALQ 435
Db 304 REKVTVMSTNPTKDLADLIGCORLPEDQKVKVQLKDLLDQILMLDPKAKRISINQALQ 363
QY 436 HPF 438
Db 364 HAF 366

RESULT 9
PCT-US04-30360-35
; Sequence 35, Application PC/TUS0430360
; GENERAL INFORMATION:
; APPLICANT: PLEXIKON, INC.
; TITLE OF INVENTION: MOLECULAR SCAFFOLDS FOR KINASE LIGAND DEVELOPMENT
; FILE REFERENCE: 039363-1703
; CURRENT APPLICATION NUMBER: PCT/US04/30360
; CURRENT FILING DATE: 2004-09-15
; PRIOR APPLICATION NUMBER: 60/503,277
; PRIOR FILING DATE: 2003-09-15
; NUMBER OF SEQ ID NOS: 167
; SOFTWARE: Patent in ver. 3.2
; SEQ ID NO 35
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Mammalian
; OTHER INFORMATION: protein sequence
PCT-US04-30360-35

Query Match 15.8%; Score 380; DB 1; Length 373;
Best Local Similarity 27.9%; Pred. No. 1.6e-17; Mismatches 126; Indels 52; Gaps 10;
Matches 96; Conservative 70; Mismatches 126; Indels 52; Gaps 10;

Qy 121 ARYEIVDTLGEAGFGKVEICIDHGMGMHVAVKVKNVGRYREARSEIQVLEHLNSTDP 180
Db 19 ARYLVRKLGWGHFSTVWLAKDM-VNTHVAMKIVRGDKVYTEAAEDIKLQRVNDADN 77
Qy 181 N-----SVFRCVQMLEWFDHGG-----HVCIVFELLGLSTYDFIKENSFLPFOIDHIROMA 231
Db 78 TKEDSGMAGNHLIKLIDHFNHKGPNVGVVVMVFELGNLALIKKYEHGPIPLIYVKQIS 137
Qy 232 YQICOSINFLHN-KLHTDLKPNILFKSDYVVKYNSKMKRDBRTLKNTDVKVDFGS 290
Db 138 KOLLGLDYMRRCGIITHTDKPNVLMETD-----SPENLIQ-----IKTADLGN 184
Qy 291 ATYDEHSTLVSTRHYRAPEVILALGWSQPCDVMSIGCILLIYVLTGPTVPTQTHD----- 345
Db 185 ACWDEHYTNSIQTWEXYSPEVLLGAPWCGGADINSTACILFELITGDFEPDEGHST 244
Qy 346 -SKEHAMMERILGPIPOHMTQKTRKYFHHNQLDWEHSSAGRYVRRCKPKL----- 399
Db 245 KDDHIAQIIEILLGSLPSYLLRNGKYTRTFNS-----RGLRNISKLKFWPLE 293
Qy 400 -----EFMLCHDEEHEKFLDLVRMLEYDPTQRTLDALQHPF 438
Db 294 DVLTEKYKFSKDEAKE-ISDFLSPLMLQDPRKPRADAGLVNHPW 336

RESULT 10
US-10-450-763-46352
; Sequence 46352, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/549,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 46352
; LENGTH: 703
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

; NAME/KEY: DOMAIN
; LOCATION: (622)..(653)
; OTHER INFORMATION: Protein kinases ATP-binding region proteins domain identified
; OTHER INFORMATION: by eMATRIX, accession number BL00107A, p-value=6.586e-11, raw
; OTHER INFORMATION: score of 18.39
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (507)..(674)
; OTHER INFORMATION: Eukaryotic protein kinase domain identified by PFam,
; OTHER INFORMATION: accession name pkinase, E-value=3.5e-27, PFam score of 101.9
US-10-450-763-46352

Query Match 14.0%; Score 337; DB 6; Length 703;
Best Local Similarity 32.7%; Pred. No. 1.7e-14; Mismatches 69; Indels 18; Gaps 2;
Matches 65; Conservative 47; Mismatches 69; Indels 18; Gaps 2;

Qy 98 KRSRSIEDDEGHILICQSGDVLRLARYEIVDTLGEAGFGKVEICIDHGMGMHVAVKVKN 157
Db 481 KFSITSFDDDEGHGFKVLVLDHIAIYRYEVLITIGKSGFGQVAKLCHKNEL-VALKIIRN 539
Qy 158 VGRYREARSEIQVLEHLNSTDPNSVFRVCVQMLEWFDHGHVCIVFELLGLSTYDFIKEN 217
Db 540 KKRPHQALMELKILEALRKDKDNTYVNVHMKDFFYFRNHFCITFELLGILNLYELMKNN 599
Qy 218 SFLPFOIDHIROMAYQICQSNFLHNNKLTHTLKPNILFKSDYVVKYNSKMKRDER 277
Db 600 NFOGFSLSIVRFTLSVLKCLQMLSEKIIHCDLKPENILVLYQKG----- 644
Qy 278 LKNTDVKVDFGSATYDDE 296
Db 645 --QASVKVIDFGSSCYEHQ 661

RESULT 11
PCT-US05-03526-8
; Sequence 8, Application PC/TUS0503526
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: MBMS AS MODIFIERS OF BRANCHING MORPHOGENESIS AND METHODS OF USE
; FILE REFERENCE: EX05-003C-PC
; CURRENT APPLICATION NUMBER: PCT/US05/03526
; CURRENT FILING DATE: 2005-01-27
; PRIOR APPLICATION NUMBER: US60/539,835
; PRIOR FILING DATE: 2004-01-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 8
; LENGTH: 1030
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-03526-8

Query Match 13.9%; Score 335.5; DB 1; Length 1030;
Best Local Similarity 26.1%; Pred. No. 3.1e-14; Mismatches 125; Indels 45; Gaps 9;
Matches 86; Conservative 74; Mismatches 125; Indels 45; Gaps 9;

Qy 116 GDVLRLARYEIVDTLGEAGFGKVEICIDHGMGMHVAVKVKNVGRYREARSEIQVLEHL 175
Db 7 GNVMN-KFEILGVVGEAGYGVVLLK-RHKETHIIVAIKKPKDSENESEVSKETTLRELQWL 64
Qy 176 NSTDPNSVFRVCVQMLEWFDHGHVCIVFELLGLSTYDFIKE--NSFLPFOIDHIROMAYQ 233
Db 65 RTLQENI---VELKEAFRRRGKLYLVFEYVEKNMELLEEMPNGVPP---EKVKSYYIQ 118
Qy 234 ICQSNFLHNNKLTHTLKPNILFKSDYVVKYNSKMKRDERLTKNVDIKVDFGSATY 293
Db 119 LKAIHWCHRNIDIVHRDIKPNLISHNDV-----LKLCDGFGFARN 159
Qy 294 ----DDEHSTLVSTRHYRAPEVILALGWSQPCDVMSIGCILLIYVLTGPTVPTQTHDSKEH 349
Db 160 LSEGNANYTEYVATRWYRSPELLLGAPYKGSVDWMSVGCILGELSDGQPLPFGSEIDQ 219
Qy 350 LAMMERILGPIQHMIOKTRKRYFHHNQLDWEHSSAGRYVRRCKPKLKFMLCHDSEH 409

Db 220 LFTIQKVLGPSEQKLFYSNPRFHGLRFPVAVHPQS---LERYLGIL-----N 267
 QY 410 EKLFVLVRMLEYDTQRTITLDEALQHFF 439
 Db 268 SVLLDLKMLKLDPADRYLTEQCLNHPTF 297

RESULT 12
 PCT-US04-30360-60
 ; Sequence 60, Application PC/TUS0430360
 ; GENERAL INFORMATION:
 ; APPLICANT: PLEXIKON, INC.
 ; TITLE OF INVENTION: MOLECULAR SCAFFOLDS FOR KINASE LIGAND DEVELOPMENT
 ; CURRENT APPLICATION NUMBER: PCT/US04/30360
 ; CURRENT FILING DATE: 2004-09-15
 ; PRIOR APPLICATION NUMBER: 60/503,277
 ; PRIOR FILING DATE: 2003-09-15
 ; NUMBER OF SEQ ID NOS: 167
 ; SOFTWARE: PatentIn Ver. 3.2
 ; SEQ ID NO 60
 ; LENGTH: 358
 ; TYPE: PRT
 ; ORGANISM: Unknown Organism
 ; FEATURE:
 ; OTHER INFORMATION: Description of Unknown Organism: Mammalian
 ; OTHER INFORMATION: protein sequence
 PCT-US04-30360-60

Query Match 13.5%; Score 326.5; DB 1; Length 358;
 Best Local Similarity 26.9%; Pred. No. 4.1e-14;
 Matches 98; Conservative 64; Mismatches 131; Indels 71; Gaps 12;
 QY 89 QSRKSHRRKRKRSSTDEEGHLCQSGDVLRYEIVDTLGEAGFKVVECIDHGMGM 148
 Db 3 QERTFYRQELNKTWEVPE-----RYQNLSPVSGAYGVCVCAAFD-TKTGL 48

QY 149 HVAVK-----IVKNVGYREARSEIOVLEHL---NSTDPNSVRCVQMLEWFDHGG 197
 Db 49 RVAVKLSRPPQSIHAKRTYR-----ELRLKMKHENVIGLDDVFTPARSLEFN--- 100
 QY 198 HVCIVFELLGLSTYDFIKENSFLPQIDHIRQAYICQSFNHLHKNLTHTDLKPENIL 257
 Db 101 DVYLVTLMGADLNNIVKQKLT---DHVQFLYQLRLGLKYIHSADI IHRDLKPSN-L 156
 QY 258 FVKSQDVVYKNSKMRDERTLKNITDKVDFGSAITYDEHHSITLVSTRHYRAPEVILA-L 316
 Db 157 AVNED-----CELKILDFGLARHTDDEMTGYVATRWYRAPEIMLNWM 198
 QY 317 GWSQPCDVWSIGCILIEYLGFTVQTHDSKEHLAMMERILGPIPOHMIQKTRKRYFHH 376
 Db 199 HYNQTVDIWSVGCIMAEILLTGRTPGTDHINQIQINRLTGTTPAYLINRMPSE----- 254
 QY 377 NQLDWEHSSAGRYVRRCK-PLKEFMLCHDEHEKFLDLVRRMLEYDPTQRTITLDEALQ 435
 Db 255 -----ARNYIQLTQMPKXNFANFVIGANPLAVDLLEKMLVLDSDKRITAAQALA 304
 QY 436 HPFF 439
 Db 305 HAYF 308

RESULT 13
 US-11-021-951-187
 ; Sequence 187, Application US/11021951
 ; GENERAL INFORMATION:
 ; APPLICANT: HAUPTS, Ulrich
 ; APPLICANT: KOLTERMANN, Andre
 ; APPLICANT: SCHEIDIG, Andreas
 ; APPLICANT: VOTSMELER, Christian
 ; APPLICANT: Kettling, Ulrich
 ; APPLICANT: COCO, Wayne Michael

; TITLE OF INVENTION: New Biological Entities And The Pharmaceutical
 ; TITLE OF INVENTION: And Diagnostic Use Thereof
 ; FILE REFERENCE: 04156.0002US
 ; CURRENT APPLICATION NUMBER: US/11/021,951
 ; CURRENT FILING DATE: 2004-12-22
 ; PRIOR APPLICATION NUMBER: 10/872,198
 ; PRIOR FILING DATE: 2004-06-18
 ; PRIOR APPLICATION NUMBER: 60/543,518
 ; PRIOR FILING DATE: 2004-02-11
 ; PRIOR APPLICATION NUMBER: 60/524,960
 ; PRIOR FILING DATE: 2003-11-25
 ; PRIOR APPLICATION NUMBER: EP 04003058
 ; PRIOR FILING DATE: 2004-02-11
 ; PRIOR APPLICATION NUMBER: EP 03025871
 ; PRIOR FILING DATE: 2003-11-11
 ; PRIOR APPLICATION NUMBER: EP 03025851
 ; PRIOR FILING DATE: 2003-11-10
 ; PRIOR APPLICATION NUMBER: EP 03013819
 ; PRIOR FILING DATE: 2003-06-18
 ; NUMBER OF SEQ ID NOS: 191
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 187
 ; LENGTH: 359
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-11-021-951-187

Query Match 13.5%; Score 326.5; DB 7; Length 359;
 Best Local Similarity 26.9%; Pred. No. 4.1e-14;
 Matches 98; Conservative 66; Mismatches 129; Indels 71; Gaps 12;
 QY 89 QSRKSHRRKRKRSSTDEEGHLCQSGDVLRYEIVDTLGEAGFKVVECIDHGMGM 148
 Db 2 QERTFYRQELNKTWEVPE-----RYQNLSPVSGAYGVCVCAAFD-TKTGL 47
 QY 149 HVAVK-----IVKNVGYREARSEIOVLEHL---NSTDPNSVRCVQMLEWFDHGG 197
 Db 48 RVAVKLSRPPQSIHAKRTYR-----ELRLKMKHENVIGLDDVFTPARSLEFN--- 99
 QY 198 HVCIVFELLGLSTYDFIKENSFLPQIDHIRQAYICQSFNHLHKNLTHTDLKPENIL 257
 Db 100 DVYLVTLMGADLNNIVKQKLT---DHVQFLYQLRLGLKYIHSADI IHRDLKPSN-L 155
 QY 258 FVKSQDVVYKNSKMRDERTLKNITDKVDFGSAITYDEHHSITLVSTRHYRAPEVILA-L 316
 Db 156 AVNED-----CELKILDFGLARHTDDEMTGYVATRWYRAPEIMLNWM 197
 QY 317 GWSQPCDVWSIGCILIEYLGFTVQTHDSKEHLAMMERILGPIPOHMIQKTRKRYFHH 376
 Db 198 HYNQTVDIWSVGCIMAEILLTGRTPGTDHIDQLKLILRLVGTGPAELLKKI----- 249
 QY 377 NQLDWEHSSAGRYVRRCK-PLKEFMLCHDEHEKFLDLVRRMLEYDPTQRTITLDEALQ 435
 Db 250 -----SSARNYIQLTQMPKXNFANFVIGANPLAVDLLEKMLVLDSDKRITAAQALA 303
 QY 436 HPFF 439
 Db 304 HAYF 307

RESULT 14
 US-10-450-763-44291
 ; Sequence 44291, Application US/10450763
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
 ; FILE REFERENCE: 790CIP3/US
 ; CURRENT APPLICATION NUMBER: US/10/450,763
 ; CURRENT FILING DATE: 2003-06-11
 ; PRIOR APPLICATION NUMBER: PCT/US01/08631
 ; PRIOR FILING DATE: 2001-03-30
 ; PRIOR APPLICATION NUMBER: 09/540,217
 ; PRIOR FILING DATE: 2000-03-31

; PRIOR APPLICATION NUMBER: 09/649,167
 ; PRIOR FILING DATE: 2000-08-23
 ; NUMBER OF SEQ ID NOS: 60736
 ; SOFTWARE: Custom
 ; SEQ ID NO 44291
 ; LENGTH: 480
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: DOMAIN
 ; LOCATION: (260)..(291)
 ; OTHER INFORMATION: Protein kinases ATP-binding region proteins domain identified
 ; OTHER INFORMATION: by eMATRIX, accession number BL00107A, p-value=4.375e-21, raw
 ; OTHER INFORMATION: score of 18.39
 ; US-10-450-763-44291

Query Match 13.5%; Score 326.5; DB 6; Length 480;
 Best Local Similarity 26.9%; Pred No. 5.5e-14;
 Matches 98; Conservative 64; Mismatches 131; Indels 71; Gaps 12;
 Qy 89 QSRKSHRRKRSIEDDEGHLCQSGDVLRLARYEIVDTLGEAGFKVVECDHGMGDM 148
 Db 123 QERPTFYRQELNKTIEWPE-----RYQNLSPVGSGAYGSCVCAAFD-TKTGL 168
 Qy 149 HVAVK-----IVKNGRYREARSEIQVLEHL---NSTDPNSVFCVQMLEWFDHGG 197
 Db 169 RVAVKLSRPFQSIHAKRTYR-----ELRLKMKHENVIGLLDVFTPARSLEEFN--- 220
 Qy 198 HVCIVFELLGLSTDFIKENSFLPQIDHIRMAYQICQSFNHLHNNKLTHTDLKPNIL 257
 Db 221 DVLVTHLMGADLNNIVKCKQLTD---DHVQFLYIQLRGLKYIHSADIHRDLKPSN-L 276
 Qy 258 FVKSDDYVVKYNSKMKRDERTLKNTDIKVDFGSAFYDDEHSTLSTVSTRHYRAPEVILA-L 316
 Db 277 AVNED-----CELKILDFGLARHTDDEMTGYVATRWYRAPEIMLNWM 318
 Qy 317 GWSQPCDWSIGCILIEYVLGFTVFTQTHDSKEHLAMMERILGPPOHMIQTRKRYFHH 376
 Db 319 HYNQTVDIWSVGCIMAEELLTGTLPFGTDLHINLQQLINRLATGTPPAYLINRMPSE----- 374
 Qy 377 NQLDWEHSSAGRYVRRCK-PLKEFMLCHDEEKLFDLVRMLEVDPTQRIITLDEALQ 435
 Db 375 -----ARNYIQSLTQMPKMFANFVIGANPLAVDLLEKMLVLDSDKRITAAQAALA 424
 Qy 436 HPFF 439
 Db 425 HAYF 428

RESULT 15
 PCT-US04-30360-17
 ; Sequence 17, Application PC/TUS0430360
 ; GENERAL INFORMATION:
 ; APPLICANT: FLEXIKON, INC.
 ; TITLE OF INVENTION: MOLECULAR SCAFFOLDS FOR KINASE LIGAND DEVELOPMENT
 ; FILE REFERENCE: 039363-1703
 ; CURRENT APPLICATION NUMBER: PCT/US04/30360
 ; CURRENT FILING DATE: 2004-09-15
 ; PRIOR APPLICATION NUMBER: 60/503,277
 ; PRIOR FILING DATE: 2003-09-15
 ; NUMBER OF SEQ ID NOS: 167
 ; SOFTWARE: PatentIn Ver. 3.2
 ; SEQ ID NO 17
 ; LENGTH: 359
 ; TYPE: PRT
 ; ORGANISM: Unknown Organism
 ; FEATURE:
 ; OTHER INFORMATION: Description of Unknown Organism: Mammalian
 ; OTHER INFORMATION: protein sequence
 ; PCT-US04-30360-17

Query Match 13.4%; Score 323.5; DB 1; Length 359;
 Best Local Similarity 26.9%; Pred. No. 6.4e-14;

Matches 98; Conservative 65; Mismatches 130; Indels 71; Gaps 12;
 Qy 89 QSRKSHRRKRSIEDDEGHLCQSGDVLRLARYEIVDTLGEAGFKVVECDHGMGDM 148
 Db 2 QERPTFYRQELNKTIEWPE-----RYQNLSPVGSGAYGSCVCAAFD-TKTGH 47
 Qy 149 HVAVK-----IVKNGRYREARSEIQVLEHL---NSTDPNSVFCVQMLEWFDHGG 197
 Db 48 RVAVKLSRPFQSIHAKRTYR-----ELRLKMKHENVIGLLDVFTPARSLEEFN--- 99
 Qy 198 HVCIVFELLGLSTDFIKENSFLPQIDHIRMAYQICQSFNHLHNNKLTHTDLKPNIL 257
 Db 100 DVLVTHLMGADLNNIVKCKQLTD---DHVQFLYIQLRGLKYIHSADIHRDLKPSN-L 155
 Qy 258 FVKSDDYVVKYNSKMKRDERTLKNTDIKVDFGSAFYDDEHSTLSTVSTRHYRAPEVILA-L 316
 Db 156 AVNED-----CELKILDFGLARHTDDEMTGYVATRWYRAPEIMLNWM 197
 Qy 317 GWSQPCDWSIGCILIEYVLGFTVFTQTHDSKEHLAMMERILGPPOHMIQTRKRYFHH 376
 Db 198 HYNQTVDIWSVGCIMAEELLTGTLPFGTDLHINLQQLINRLATGTPPAYLINRMPSE----- 249
 Qy 377 NQLDWEHSSAGRYVRRCK-PLKEFMLCHDEEKLFDLVRMLEVDPTQRIITLDEALQ 435
 Db 250 -----SSESARNYIQSLTQMPKMFANFVIGANPLAVDLLEKMLVLDSDKRITAAQAALA 303
 Qy 436 HPFF 439
 Db 304 HAYF 307

Search completed: March 13, 2005, 00:18:18
 Job time : 30 secs

This Page Blank (uspto)